A Corpus-based Functional Analysis of Music Discourse

Comparing Expert and Novice Academic Writing

Thesis submitted in partial fulfillment

of the Ph.D. in Applied Linguistics

Submitted: April 22, 2024

Justin Berg

School of Education and English

Word Count: 74,073

Abstract

Music is central to the human experience. While its signification may be shaped by language, one of the primary linguistic discourses about musical meaning, that of music scholars, is itself delimited by another semiotic: musical score notation. Applied Linguistics research has sought to illuminate the structures of disciplinary-specific discourses in English, the lingua franca of academia, but has yet to investigate the discourse of music scholars. These scholars work within the field of Musicology and related areas of Ethnomusicology and Music Theory. The present study extracts lexical bundles from two purpose-built corpora of expert and novice writings. These bundles are then categorized and analyzed according to structure, function, and content. The interactions of these bundles with ambient propositional content are then examined and categorized to understand how this interaction constitutes patterns unique to this specific disciplinary discourse at both the expert and novice level. From this, the epistemological concerns of the discipline are extrapolated. Findings from both corpora are compared to determine the degree of competence exhibited by novice writers.

Contents

Table of Data Graphics

Acknowledgements

1 Introduction

1.1 Genesis

The idea for this research arose from my own experience of disciplinary writing classes in Music Performance and Musicology degree programs. Those courses largely focused on sources and citation rather than writing. Without specific guidance on how to structure an argument and what grammatical choices were available for that purpose, I was ill-equipped to analyze and improve my own writing; nor did the feedback I received address this problem. In the years since graduation, I have continued to wonder how academic music writing works; hence, this research.

1.2 Language for the Ineffable

As music possesses an ineffability all its own, it poses a challenge for listeners who want to share their experience of certain pieces or performances with others. When discussing it, speakers must fall back on metaphors, analogies, inexact comparisons; but as one of my mentors liked to say, 'every example limps and every metaphor is way off base by design' (Rev. William Padavick, personal communication). It is this imprecision of expression that entices humans to write about music. The act of bringing language to bear on its ineffability represents a conviction that it is possible to elucidate music's meaning. As a semiotic, music has the potential to

communicate, yet there may be less agreement among listeners as to the meaning of pieces heard than can be achieved through words. Hence, language is employed as a crucial means of illuminating musical meaning. To achieve this feat, detailed explanations of music are most often conducted through the medium of writing. At the academic level, this is accomplished through writing and publication. To understand why language is considered an indispensable means of understanding music, it is necessary to examine the nature of Music-as-phenomenon and how that phenomenon evokes the otherworldly (Tuan 2009).

Before proceeding to the discussion of writing about music, it is first necessary to consider the nature of music itself to better understand why writing is applied to it at all. In essence, this is a necessary procedure precisely because music is so notoriously difficult to define. In its simplest form, this difficulty is observed in disagreements between listeners as to what constitutes music, such as the debate regarding the musicality of indeterminacy. For instance, does John Cage's *4'33"* constitute a piece of music, given that the performer is never instructed to touch a single key of the piano? In its most complex form, this same difficulty is observed when trying to define the nature of more conventional pieces. For example, what constitutes the work known as Bach's *St. Matthew*

Passion? Is the work somehow the totality of all actualized and potential performances, as well as its various score iterations, or is it something other or more than that? To answer such philosophically intricate questions, it is necessary to write, for through writing ideas are developed and debated. Thus, the following section considers the nature of music from a philosophical perspective in order to better situate it as a phenomenon to be explored through the written word.

1.3 The Nature of Music

Music offers a unique means of representing human experience. No doubt this explains why there has never been a society devoid of it (Brown 1991; Savage et al. 2015; Eegermann et al. 2015). Though music is imperative, '[w]hat music is remains open to question at all times and in all places' (Bohlman 1999:17). Though we experience it as an acoustic event in daily life, music also exists beyond any single manifestation of itself in time; no performance or notation can be said to represent the totality of the phenomenon any more than a sole hearing can reveal the entire substance of a piece (Gadamer 2004b; Kivy 1983, 1987). Understanding music, then, requires not only comparison of acoustic events but also a means of surveying the phenomenon, a gallery from which to listen.

Despite differing ontologies, philosophers of music agree that temporality is intrinsic to the musical experience (Matheson & Caplan 2011). Nonetheless, these same ontologies of music must address the dichotomy of music manifest in and beyond time by considering the nature of the musical experience (manifest in time) and that of the musical work (existent through time) (Kania 2017; Scruton 2016).[1] Musical experiences are auditory events; musical works are discrete, notated pieces. Though an oversimplification, herein it will suffice to speak of musical experiences as performances and works as compositions. Performance is identifiable as a configuration of place, time and participants, and is further distinguished by performance practices, mistakes and listeners' reactions, among other factors. Such a musical experience is a unique occurrence, not equivalent to any work. As a composition, a work — including its place of origin, composer, form of notation, history of performance and reception — is an entity larger than any given time or place (Goehr 1992). Thus, experiences of a work manifest at a single time, while the work itself exists through time (Young 2011; Rohrbaugh 2003).

Because of this intra-temporal nature, musical works can be objects of investigation beyond their manifestations. Of course, not all musical experiences are actualizations of musical works. Indeed,

some cultures neither employ notation nor have a notion of *the work,* yet musical experiences not definable as works may resemble each other to the degree that they can be classified, by form or function, and thus presented for consideration beyond their manifestations. Given their temporality, all musical experiences share the essential property of non-duplicability (Matheson & Caplan 2011; Gadamer 2004b). The sheer number of variables in any performance insures that, even were it physically possible to interchange performances, aspects of the experience would be altered. As a consequence, anyone seeking to understand the meaning of a category or work, even musical materials such as rhythm and pitch, must contemplate them as existent through time, rather than as individual manifestations (Davies, S. 2011). If such contemplation is to be shared, the social semiotic of language (Halliday 2001) is needed to enable dissemination of that understanding across time and space (Kramer 2003).

1.4 Discourse about Music

On account of its abstract nature, musical meaning depends on context (Bohlman 1999), and explication of music in context can only be accomplished through words (Kivy 1983, 1987). While such discussion could be verbal, writing is more advantageous for such philosophical consideration as it renders detailed analysis more

permanent. Since writing can be developed and revised over time, it is especially conducive to thorough examination of musical works and categories (music beyond time), and experiences (music in time). Insomuch as writing boasts ample space for ideas to grow and play, it is uniquely capable not only of describing musical meaning but moreover of *reconstituting* it (Kramer 2003:128) by contextualizing it. Because music — at least that music which does not rely upon language, such as purely instrumental music — lacks clearly discernible external signification (Gadamer 2004b; Scruton 2016; Tuan 2009), it depends on this reconstitutive property for signification (Bowman 1998; Nattiez 1990); hence, the value of writing about music.

Partially as a result of this dependence, music is open to varied interpretations (Kania 2017). Writing not only represents music (for it is more than designation), it re-presents[2] it, making present again not only the musical object but moreover its initial context for consideration. Writing may then intercede on behalf of music for the purpose of presenting forgotten context or novel interpretations to the reader. By asserting new interpretations, writing transforms music, rendering it newly meaningful (Gadamer 2004b; Kramer 2003). Analytical writing may present music in ways that cannot even be acoustically experienced by listeners (Zbikowski 1998). In

thus making music present again, writing enables readers to step beyond the temporality of experience to contemplate music as it exists through time; it grants access to what Nattiez (1990:42) terms the *total social fact* that is music.

This anamnestic quality wherein writing re-presents music, is immensely valuable for musical description because music is 'fundamentally ineffable' (Kramer 2012:101). 'Music relies neither on linguistic order nor on physical context, but on organization that can be perceived in sound itself, without reference to context or to semantic conventions' (Scruton 2016:5). Writing, then, is a potent resource for conferring meaning on the perceived structures heard in music. Thus, it can serve as a hermeneutic of music, whereby it may wield enormous influence over subsequent performance and reception. Indeed, successive applications of this hermeneutic have established it as a prominent re-creative act that is indispensable to an understanding of music as phenomenon and instance. This is especially apparent in the work of music scholars and performers, who depend on writing as their primary means of communicating with their peers.

Writings on music have always sought to influence performance, pedagogy and analysis. This has been evident in English for some

five centuries, beginning with the pedagogical writings of the Landsdowne Manuscript (Lansdowne MS. 763, c 1200-c 1700). Over time, English discourse on music grew to meet the needs of an expanding musical scholarship. The present investigation focuses on writings of the current scholarly community, both experts who present their interpretative arguments about music through writing and novices who aspire to the same level of expertise, particularly those for whom English is a second language. Though these students account for a significant percentage of learners in the global academy, their use of English remains under-explored (Mauranen et al. 2010).

Given its status as the *currency of intellect* (Pratt 2011:ix), writing is both a vital practice and a critical skill (Kellogg & Raulerson 2007) which musicians are compelled to develop if they wish to communicate formally about their subject, let alone participate in the expert discourse of the scholarly community. Despite this fact, no linguistically informed analysis of such music writing has been undertaken previously. Perhaps the abundant linguistics research on science, medicine and finance, among others (Flowerdew, L. 2002), helps explain this lacuna, as would a focus on language-based arts such as literature. Whatever the case, any variety of academic discourse requires a common ground, a standard, by which

knowledge may be disseminated (Swales 1990; Johns 2003; Gee 2014b). Without such an understanding, academics risk missing or disregarding the musical trends that are an outgrowth and reflection of, as well as a commentary on, human experience; what is at stake is nothing short of understanding who we are.

Despite the significance of this discourse, music scholars are seldom equipped to offer a linguistic perspective on their own disciplinary writing, a situation common to academics in various fields (Gebhard et al. 2013). A few investigations have been conducted into other discourses of music, such as Pérez-Sobrino and Julich (2014), who found that verbal descriptions of music employ metaphor to convey meaning. Some musicologists have also taken up questions about the intersections of music and language, though not from a linguistic perspective. Among these, one of the most interesting and controversial theses is that of French music semiologist Jean-Jacques Nattiez (1987), who argues that discourse about music is inherently analytical and should therefore generate a unified discipline of Musicology. As this research will show, however, the present form of music discourse is more akin to description than analysis; nor is this form the only conceivable one. Rather, the academic discourse of music has developed into this descriptive form due to a continued focus on the canon of Western music and

the constraints on the discourse imposed by reliance on score notation.

1.5 Personal Experience

In my own studies several years ago, all music students were required to enroll in writing courses at each stage of undergraduate and graduate study (N.B.: In North American tertiary education, 'graduate' is the equivalent label for the British system's 'post graduate' level). These were taught by a professor of music during the first year of study for each degree. The stated purpose of these courses was to teach students the form of disciplinary writing and thus initiate them into the community of music academics. Faigley and Hansen (1985) note that such discipline-specific writing courses are a typical requirement of first-year university study.

In the three writing courses I studied, instruction focused almost exclusively on research, citation and style. Regarding this last, we novices were admonished to write with accuracy and clarity (for which no explicit guidance was given), as judged by grammar and intuition, respectively. As required texts for these courses were guides to style and citation, it became apparent that structure and syntax would not be addressed, except by way of editorial correction. Indeed, the expert instructors were not entirely prepared

to give constructive feedback beyond noting prescriptive grammatical rules or punctuation errors.

As a student, I was unsatisfied with this kind of instruction and wanted to know more about the mechanics of writing: how to make decisions regarding organization and structure, how to methodically produce coherent writing, and how to employ grammatical constructions appropriately. Of course, my grammatical knowledge was based on a conventional parts-of-speech view, but that did not aid me in the construction of texts from groups of words to larger, cohesive stretches. From my current perspective as a student of linguistics and a writing tutor, it seems that my own student writing could have benefitted greatly from explicit instruction on issues such as structure, cohesion, stance and hedging, to name but a few points. Instead of such a functional view, however, the courses I studied were oriented toward a prescriptive view that employed conventional style guides, which provided little instruction about grammar and none whatsoever regarding the larger concerns of textual organization. Presently, there is a profusion of such guides, demonstrating that they remain popular for music writing instruction.

1.6 Music Writing Guides

In recent years, the type of style guide employed as writing textbooks for music courses has proliferated. Currently available titles include: Poultney (1995), Irvine and Radice (2003), Bellman (2006), Wingell (2008), Herbert (2009), Holoman (2014). All six of these books include the word *Writing* in their titles, which speaks to the prestige of, and demand for, disciplinary writing in music. These guides advocate accuracy, clarity and attention to purpose. The following excerpt illustrates how such guides advise novice writers:

> The basic points to emphasize about content, form, and style are:
> - Your writing must be fit for its purpose: its content, form, and style must serve the needs of your intended readership.
> - Irrespective of those aspects of style that distinguish your writing from that of others, your writing must be clear; it must be so clear that everyone who reads it takes from it only the meaning that you wish to convey.
> - What you write must be founded on a clear and accurate understanding of your subject (Herbert 2009:13-14).

There is not much here from which to learn, perhaps aside from the concern to write purposefully and engage one's audience. Similarly, Wingell's (2008, Fourth Edition) volume, *Writing about Music*, does little more than decry the state of undergraduate writing (not wholly an original complaint), admonish students to avoid subjectivity, and offer checklists of grammar, terminology and citation. There is irony in such exhortations to lucid prose when the same authority does not elucidate how to produce them. Other guides offer less help, such as Holoman's (2014) tepid suggestion to consult university writing websites for assistance producing purposeful prose. These books epitomize the notion held by many academics that their formal discourses are tautly organized, while they themselves remain uncertain how to explain such organization, leaving description to the linguist (Sinclair 2004b). Even then, most experts disregard the assumption that their disciplinary discourse possesses such a snug structure (ibid.), which may not be the case at all, or at least not in the sense it is assumed. Obviously, such books leave a great deal to be rendered in finer detail regarding academic writings about music, particularly concerning the interface of structure and content. Presumably, these writers esteem the overall communicative goals of such writing, yet throughout these books there is barely any acknowledgement, beyond admonishments regarding topic sentences, of the structures that bind texts together

to elaborate meaning. Surely these authors assume the importance of writing competently; nonetheless, they remain unequipped to discuss the specific disciplinary standards and writing structures that constitute this unique discourse about music.

Among the music professors who taught my writing courses, one was an established performer who wrote his own program notes and two were published academics. All three wrote engaging prose, having had years of practice. The knowledge they lacked about writing was not disciplinary, of course, but linguistic. In fact, educators typically do not possess a detailed understanding of the disciplinary discourses they themselves read and write, and therefore are neither able to adequately indoctrinate students in the same, nor to demonstrate how such discourses forge disciplinary-specific meanings (Gebhard et al. 2013). Similarly, musicologists are not likely to be aware of the ways in which their discourse is constructed linguistically. As a result, subject experts tasked with writing instruction seldom specify expectations for language acquisition pertinent to their field, which leaves novices to flounder (Hyland 2006). Given that writing competence develops under the influence of disciplinary learning (North 2005; Gimenez 2011; Qin 2014), this is a particularly lamentable situation because it means that students' education remains incomplete. Thus, the expert

community not only fails to induct new members as fully and effectively as possible, but even inadvertently bars their admission to the community.

Several years after completing music studies, I moved to China, inaugurating a second career in English Language Teaching (ELT).[3] Being placed in the position of expert instructor afforded me a new appreciation for the challenges of academic writing, especially for L2 novices (i.e., those learning in a second or additional language). Indeed, the majority of my students did not even fit Gebhard et al.'s (2013:107) profile of second language learners who possess 'the semiotic resources required to construct everyday meanings in a second language, but…struggle to construct discipline-specific meanings despite years of schooling'. The above-mentioned writing guides offer no assistance for framing propositional content and organizing it according to both the norms of the academic register and those of the discipline of Music. Without explicit instruction in these areas, novices are left to achieve a degree of expertise through trial and error, engaging with feedback as best they are able. This lack of informed guidance no doubt renders the writing process more onerous than it initially seems, which surely discourages novices from making meaningful, substantial progress in this critical skill.

Following music studies and subsequent years of work as a performer and teacher, I returned to university as a master's student in Applied Linguistics to research language and learning needs for high school music students (ages 14-18) in International Baccalaureate (IB) Programmes throughout China. I found that the students who participated in my research (Berg 2015) aspired to a greater ability to express their musical experiences and ideas, which affirms Slabakova's (2014) observation that L2 learners naturally want to express themselves in the target language.[4] Meanwhile, their music instructors bemoaned the state of their students' writing abilities (Berg 2015). This experience corroborated Gebhard et al.'s (2013) statement that, while younger learners may be exposed to certain conventions of writing, they are increasingly unlikely to be schooled in more complex forms as they process through higher grades, where academic language is typically relegated to a position that is secondary to subject courses. Given that the teachers of those courses are subject experts, they are not expected to have specialized linguistic knowledge. This being the case, both instructors and students could benefit substantially from a more detailed understanding of the inner workings of music discourse.

Given my background as a music student and my experience conducting linguistic research among music students, I chose to

focus the present investigation on academic music writing. Because this discourse has not been previously researched, I looked for a method that would help me gain an overview of it, one that would permit the creation of a map of the content and most frequent features unique to this discourse. This seemed appropriate for an initial investigation into this uncharted territory since any more specific approach would have to sacrifice a global view for a granular one that magnifies select aspects.

1.7 Lexical Bundles

To balance macro and micro perspectives, a point of entry to this discourse was needed that was frequent and thus pervasive, but that would also permit examination of the propositional content connected to such frequent items in order to discern patterns of such content across the discourse. This resulted in a focus on the intersection of invention and formulaicity within the discourse, which necessitated a quantitative approach based on a large amount of data. The most obvious solution was to employ a corpus-based methodology. As corpus is largely focused on, and useful for, investigations at the lexical level, I sought avenues of corpus investigation that extend beyond the collection of individual words or even collocations. Since corpus facilitates searches for frequently recurring items, I chose lexical bundles as my point of departure, as

these strings of words constitute the longest textual features to be mined from corpora by frequency (Biber et al. 1999).

This seemed an interesting approach to the problem of mapping music discourse given that lexical bundles have been researched within academic writing and elsewhere, but have not been widely used to investigate content within a given discourse. With that in mind, I built dedicated corpora and mined them for lexical bundles to gather and analyze the ambient propositional content on either side of these bundles, using a systemic functional view of groups and phrases. This approach permitted investigation of high-frequency formulaic items and propositional content, the original or innovative dimension used to build this discourse. This type of methodology required a significant amount of manual analysis of the ambient content, as no technology is perfectly capable of extracting propositional content from a corpus according to a set of systemic functional parameters.

Though previous research into lexical bundles in academic writing has yielded lists of such bundles common to this register, such research typically focuses on bundles of four or more words. This is partly due to the fact that 2- or 3-word bundles are so plentiful as to be impractical for limited study. Nevertheless, building corpora of

limited size opened the way to the unusual approach of mining for 3-, 4- and 5-word bundles, taken for two reasons. The first was the idea that 3-word bundles, given their higher frequency by a factor of 10 (compared to 4-word bundles), might yield examples containing disciplinary-specific terms, a supposition that proved correct. The second reason was the thought that as much information as possible should be extracted from both corpora in order to map this discourse as thoroughly as possible. Since no such research had been conducted previously, I was unsure what might appear in a corpus of music discourse texts, and therefore adopted an exploratory approach to the research, rather than a hypothesized one. The result is a wealth of lexical bundles and ambient content that has provided an enormous amount of data for mapping this discourse.

1.8 Research Aims

In light of the universality of music, its social value, the power of writing to re-present and transform musical meaning, the disciplinary demand for such writing, the dearth of subject specialists' knowledge regarding their own discourse, and the lack of linguistic research into this discourse, an analysis of music discourse is surely overdue. While differing forms of this discourse could be analyzed, the present research focuses on academic writing as the

preeminent form of that discourse employed throughout the expert community for the purpose of disseminating knowledge. Thus, this project is undertaken for music experts and novices alike, whose scholarly economy is dependent upon the common currency of writing, and whose mandate it is to know the world.

Discourse analysis may begin from various perspectives, ranging from the study of language to the social practices surrounding it (Tannen et al. 2015), and given the lack of research into music discourse, any of these aspects could serve as a point of departure. As the aim of this study is to assist novices with their writing, however, the text itself has been chosen as the most imminently useful facet to illuminate. In particular, this thesis considers the intersection of formulaicity and originality by investigating how writers, expert and novice, employ formulaic language to present propositional content.

Inspired by Biber et al.'s (2004:376) *deliberately exploratory* approach, this project mines two specialized corpora for frequent formulaic items, herein referred to as lexical bundles (hereafter LB). These are then categorized and analyzed both structurally and functionally to determine how these items connect with their ambient propositional content to project the epistemological

concerns of music discourse. Since one of the greatest challenges for novice writers is the production of typical structure (Wray 1999; Green et al. 2000; Wray & Perkins 2000; Johns 2003; Cortes 2004; Chen & Baker 2010; Winberg et al. 2010; Mollet et al. 2011; Crossley et al. 2012; Staples et al. 2013; Alhassan & Wood 2015; Ebeling & Hasselgård 2015; Chen & Baker 2016; Crossley et al. 2016), it is hoped that this approach will afford a broad perspective of disciplinary epistemology through a detailed investigation of formulaicity and structure. A view into the inner workings of this discourse could aid in bridging the *differential experience* that distances novices from experts (Vickers 2010:116). Because students learn disciplinary writing best within subject courses (North 2005; Gimenez 2011; Qin 2014), it is vital that experts enhance their teaching of this discourse with precise linguistic description.

1.9 Terms and Conventions

Before proceeding, it will be useful to make a distinction about the subsequent use of several terms, most notably *Music Discourse*. As the present research proposes to investigate discourse *about* music, rather than music *as* discourse, the majuscule *M* will be used throughout this thesis as an orthographic aid specifically referencing the academic discipline of Music, and disambiguating the same from other senses of the phrase, including but not limited to phenomena

of music, such as performance or the experience of listening to music. As a further convention, *Music Discourse* will herein specifically denote writing originating from any sub-field within the discipline of Musicology, most especially Ethnomusicology and Music Theory. Table 1.9.1 lists all of the abbreviations and terms used herein with their particular orthographic conventions.

Table 1.9.1: Acronyms and Terms

Acronym/Term	Meaning
CL	Corpus Linguistics
EAP	English for Academic Purposes
ELT	English Language Teaching
ESP	English for Specific Purposes
ExCo	Expert Corpus
L1	First Language
L2	Second (or additional) Language
LB	Lexical Bundle(s)
Music Discourse	Musicological Writings
NoCo	Novice Corpus
SFL	Systemic Functional Linguistics
SLA	Second Language Acquisition

2 Literature Review

2.1 Outline

The literature review begins with an explanation of discourse, the central topic of this thesis. Following this, the notion of discourse community is discussed as it relates to expert members and novices. Then, a brief history of relevant areas from Applied Linguistics is given, culminating in a discussion of English for Academic Purposes and the concept of register. Lastly, the two key areas of the present methodology, Corpus Linguistics and Discourse Analysis, are considered, the latter of which returns to the opening discussion about discourse.

2.2 Discourse and Discourse Community

The primary goal of this research is to map the discourse of academic writing about Music (i.e., Music Discourse). Thus, it is first necessary to contextualize the concept of discourse. Gee (2014b:148) enumerates discourse thus:

> The key to Discourses is "recognition." If you put language, action, interaction, values, beliefs, symbols, objects, tools, and places together in such a way that others *recognize* you as a particular type of who (identity) engaged in a

particular type of what (activity), here-and-now, then you have pulled off a Discourse.

Consequently, recognition by others of those engaged in a discourse is critical. Such recognition is only possible when everyone involved shares similar knowledge and experience, as enumerated above. Though sharing a perspective, it is unlikely that those able to recognize a discourse would feel compelled to define it exactly, either for each other or for those unable to recognize it, inasmuch as the ability to use a discourse and be recognized for doing so would seem sufficient to most practitioners. This is Gee's (ibid.) sense of having 'pulled off a Discourse'. Thus, discourse tends to be a fundamentally opaque, communal phenomenon, the inner workings of which remain hidden from all except those few who would attempt to analyze it.

This social dimension of discourse was illuminated by the constellation of mid-20th century intellectual movements constituting the *social turn.* This reorientation of philosophical thought, away from individual and behaviorist theories toward social constructions and relationships, situates discourse as a social phenomenon (Gee 1999), as in Paul-Michel Foucault's proposed paradigm of discourse as 'language-in-action' (Blommaert 2005:2;

see Foucault 2002a, 2002b). This view of discourse as socially situated and constructed by choice is also fundamental to Halliday's (2001) systemic functional description of language as social semiotic. Due to this social dimension, context is understood as paramount because it is 'the totality of conditions under which discourse is being produced, circulated and interpreted' (Blommaert 2005:251). Much research has focused on the social context of discourse, such as Escudero's (2011) findings that patterns of language use reflect institutional identity. English and Marr (2015) apply this to educational institutions, noting that both academia and its various disciplines constitute discourse communities that arise from common practice and experience. As demonstrated in Chapters 4 and 5, this is also true of Music Discourse.

From these top-down perspectives, communities of language users are understood as coalescing around a shared purpose while also forming conventions that govern language use for that purpose. Thus, Peregrin describes language as *proprietary*, in the sense that communities maintain agreed standards of correctness in language production (Peregrin 2012:210), a linguistic phenomenon familiar to anyone who has ever observed a group of people with their own 'in speak'. Swales (1990:25-27) details this phenomenon of shared, specialized norms by contextualizing it within the notion of the

discourse community: a socio-linguistic entity that 1) broadly agrees on common goals, 2) possesses a means of communication between members, 3) uses these means to give and receive information, 4) has one or several specialized communicative genres, 5) possesses specialized lexis, and 6) maintains 'a threshold level of members' who possess expert knowledge of content and discourse.

Applying Swales' criteria to the discipline of music, it is evident that such a discourse community exists for music. The principal academic disciplines comprising this community are Ethnomusicology, Musicology, and Music Theory. These communities consist of scholars, though also some performers, who possess or are pursuing academic degrees in the study of music. Herein, these disciplines together are simply referred to as *Musicology*, the community (denoted herein by the majuscule *Community*) of which consists of scholars engaged in descriptive, historical, theoretical or social analysis of the phenomenon of music. These scholars meet all of Swales' criteria for a discourse community since Musicology possesses: 1) shared goals of investigation and research into Music; 2 & 3) various publications and gatherings that facilitate communication and exchange of information (as demonstrated by the journal articles constituting the Expert Corpus of this study); 4 & 5) specialized multimodal genres, published or recorded, that

employ a wealth of disciplinary lexis, as evidenced by the existence of dictionaries devoted to the same; and 6) a membership threshold, typically in the form of terminal degrees, publication, and academic posts.

Merging Swales' focus on communication with Gee's notion of recognition, Hyland (2009:60) defines academic discourse communities thus:

> International currency is a key criterion, particularly the extent to which leading universities recognize the independence of an area and give it the status of a department with professorial chairs, budgets, and degrees; whether a distinct international community has appeared around it with the professional paraphernalia of conferences, learned societies, and specialist journals; and whether the wider international community generally perceives it to have academic credibility and intellectual substance.

Thus, recognition and specialized communication are central to the creation and sustenance of the Music Discourse community. Indeed,

Musicology meets all of Hyland's criteria, being internationalized, recognized and replete with all the apparatus mentioned above.

2.3 Discoure Experts

One corollary of communal recognition is that members of a discourse community act as gate-keepers, admitting only those recognized as experts (Gee 2014b:148). Generally, this function serves to initiate novices into expertise. For instance, the larger of the two corpora in this study is constructed from texts written by Musicologists, the gate-keepers for novices seeking admittance to the Music Discourse community. These novices are typically learners pursuing music studies for future careers as performers or scholars. The second corpus in this study is comprised of essays from such novices.

Although Music Discourse is a product of the Musicology Community, the writings constituting it are read by experts and novices in the areas of Musicology and Performance. Academically, the distinction between Performance and Musicology studies is institutionalized in differing postgraduate degrees, the Doctor of Musical Arts (DMA) representing the former and the Ph.D. the latter.

Experts in Musicology and Performance are members of the Discourse Community and producers of its Discourse. Their inclusion signifies that they have successfully completed an apprenticeship of sorts, having invested time and energy in study, including a substantial writing component. This fact, however, is by no means a guarantee of their expertise as writers. The mere ability to write prose that is accepted by a discourse community is not, ipso facto, evidence of good writing practice, nor of an ability to teach such writing to others. Rather, this ability signals that a writer has been permitted to pass through the communal gate by adopting its linguistic proclivities while acceding to its critical and editorial demands.

It is also entirely possible that experts possess the ability to write exemplary disciplinary prose without precise grammatical and syntactic knowledge of how they themselves accomplish this, which is to say that they may be expert users of their own language for the purposes of their own discipline without possessing a grammarian's understanding of that expertise. Regardless of individual writing ability, experts not only regulate discourse conventions, but also participate in shaping their discourse through continued production and consumption of it. The configurations of features found in such texts establish the discourse norms that

novices must master to gain admittance to the community. Since some of these features are recurrent and others not, the intersection of formulaicity and originality has been chosen as a central concern of the present investigation.

Generally, musicologists are the academic experts constituting the majority of the Music Discourse Community, though some performers choose to be actively involved as well. Musicologists situate their scholarly identity within a lengthy history of music scholarship, perhaps in part to legitimize Musicology's rather brief history. Musicology began in the mid-nineteenth century as an attempt to describe and contextualize music in the human experience. From its beginnings in Germany as *Musikwissenschaft* (Duckles et al. 2001), the discipline has been largely conducted through the production of specialized writing, in journals or other publications. As in any discipline, musicologists promulgate their expertise through these writings; they 'own' and 'operate' the Discourse (Gee 1992:107). As a consequence, novices must comprehend both the content and linguistic forms of the Discourse to earn 'credentials as an insider' (Gee 2014b:147); that is, to earn recognition. Hyland (2006:3) summarizes the challenge novices face: 'There is now compelling evidence across the academic spectrum that disciplines present characteristic and changing forms

of communication which students must learn to master in order to succeed'. Considering the highly specific challenge this presents, it would benefit novices if experts could elucidate the forms and mechanics of such communication for them, thus further enabling their bid for admittance to the Discourse Community. Given the evaluative role played by experts (Bhatia 2004), instructors would likewise benefit from a modicum of linguistic training in their discourses so as to better assist novice writers with their communal induction (Kennedy 1983).

Since writing is both a requirement and a benchmark for novices striving toward induction into the Community, as demonstrated by its privileged position academically, its mastery poses a major and fundamental challenge to novices. Because Music Discourse has not previously been analyzed, it is hoped that findings from this research will be of value to experts tasked with writing instruction. Given the many requirements of academic writing, such findings could prove particularly useful to novices whose first language is not English, since they must meet the demands of general language learning as well as those specific to a given type of writing, while simultaneously mastering disciplinary content.

2.4 Second-Language Acquisition

To analyze and compare the writings of music experts and novices, I have drawn upon linguistic theories and methods developed within the field of Applied Linguistics. This field is concerned with theoretically informed solutions to the problems of language in the real world (Simpson 2011), or in Brumfit's well-circulated definition: 'The theoretical and empirical investigation of real-world problems in which language is a central issue' (Brumfit 1995:27). Because of its emphasis on the application of theory to practice, Applied Linguistics has contributed significantly to the field of language learning. Avenues of research focused specifically on second-language learning and disciplinary communities, which are relevant to an analysis of Music Discourse, are the topic of this section.

Following World War II, Britain and America promoted English Language Teaching (ELT) globally (Howatt & Widdowson 2004; Kaplan 2010). A complex endeavor, ELT grew to encompass both teaching and research activities intended to address the needs of English learners from any background (Hall 2016). As ELT spread, its classrooms filled with students from around the globe. Though some distinguish these students as *non-native speakers*, such labels are problematic as cultural reincarnations of imperialism (Canagarajah 1999; Phillipson 1992, 2012; Llurda 2016); thus,

these labels have been exchanged for more neutral ones. One such example is the pair of designations L1 and L2 (Slabakova 2016; Dewaele 2018), denoting not the speaker, but rather the speaker's first language and second or additional language, respectively (Johnson & Johnson 2001; Ortega 2011). Though these labels could still be criticized as insufficiently neutral, they are employed herein as an expedient means of distinguishing between first- and second-language learners. Most often, however, the term *novice* is used herein to refer both to learners generally and L2 learners specifically.

In recent decades, the number of L2 English learners has swollen to a population some three to fives times larger than that of L1 students (Doughty & Long 2003). Naturally, L2 learners 'would like to understand and to be able to convey thoughts and feelings and observations in another language' as they do in their L1 (Slabakova 2014:127). For these reasons, the learning experiences of L2 students prompted the development of Second Language Acquisition (SLA). Begun in the 1960s, SLA continues to focus on the ways in which people learn or acquire an L2, particularly in a context where the language is in general use (Ellis, R. 2015). Because of its concern with the application of language learning in daily contexts, the scope of SLA encompasses learner materials, teaching

applications and even special learning needs (Doughty & Long 2003:3). The primary goal of SLA research has been 'to characterize learners' underlying knowledge of the L2, i.e. to describe and explain their COMPETENCE' (Ellis, R. 2008:6). As a result, SLA has come to embrace a broad perspective known as *communicative competence,* which entails understanding how to use L2 grammar in real-life situations (ibid.; Block 2003). In addition to grammar, the descriptor *competence* has been applied to other aspects of language learning, including discourse (Bruce 2008).

2.5 ESP to EAP

As noted above, the notion of competence can range beyond a novice's command of grammar to include linguistic facility with a given domain of knowledge (Bruce 2008; North 2005; Gimenez 2011; Qin 2014). Recognition that such domains are integral to language learning led to the development of a new field within Applied Linguistics, English for Specific Purposes. ESP is sometimes contrasted with general English in that ESP learners focus their studies on language specific to the communicative demands of a particular skill or domain of knowledge, rather than on daily usage (Hutchinson & Waters 1987; Dudley-Evans & St. John 1998; Belcher 2009; Johns 2013; Starfield 2016). As such, ESP forms part of the

professionalization process for L2 learners seeking employment in English-speaking institutions.

Because of its orientation towards purpose, ESP draws on authentic material as its primary learning resource (Carver 1983:133). Since its inception in the early 1960s (Johns 2013), ESP has been embraced in numerous places globally (Hall 2016). Its encounter with other areas of linguistic study, particularly Halliday's Systemic Functional Linguistics, adjusted ESP's earlier trajectory toward the analysis of discourse and register (Starfield 2016). Essentially, register is defined as 'recognizable configurations of linguistic resources in certain texts' (Thompson 2014:40), such that texts sharing a register may be recognized as having a common purpose. Academic writing constitutes one type of register since examples of such prose exhibit a particular configuration of shared features, cumulatively known as the academic register. Because register is connected to domains of knowledge and specific discourses, Halliday et al. (1964) suggested a top-down view of these domains and discourses by proposing that register be studied in order to address the particular problems of teaching English to professionals. As ESP instructors were interested in patterns of recurrent features found in academic writing, ESP research adopted register analysis to prioritize high frequency features in the teaching of domain-specific

content (Hutchinson & Waters 1987). Though recurrent features within the academic register have been generally well-researched, not all disciplines within that register have been analyzed equally thoroughly. For instance, this study appears to be the first to analyze high-frequency multi-word features along with their ambient propositional content in Music Discourse.

As noted with the top-down view proposed by Halliday et al. (1964), the study of configurations of recurrent features provides an overview of discourse from the perspective of frequency and recurrence. An approach of this type inherently prioritizes patterns of language use. This enables the analyst to see language from a global perspective, rather than from a granular one departing from lexis. As ESP began to investigate registerial patterns, the emphasis on frequency caught the attention of corpus linguists, whose form of investigation is inherently lexical in nature. The result of this has been an intersection of registerial and lexical research (Paltridge 2013). Inspired by this intersection, I have used corpora to facilitate a qualitative examination of disciplinary content connected to formulaic sequences in Music Discourse that is based on substantial quantities of data.

Over the past half century, the cross-fertilization of ESP and Corpus Linguistics has resulted in a proliferation of ESP specializations, ranging from science, to medicine, to finance (Paltridge 2013; Flowerdew, L. 2002; Atkinson 1999), though rather less research seems to have been conducted in the Arts and Humanities. The access to authentic materials offered by Corpus has aided ESP by enriching pedagogical materials — identifying disciplinary lexis (Basturkmen 2010), for example — and assisting language instructors to learn disciplinary content (Starfield 2016), contributions that offer obvious benefits for novices and experts alike. As there are still unexplored areas of specialization, the disciplinary focus of ESP, coupled with a corpus-based methodology, continues to offer productive avenues for research.

As ideas from ESP, including register analysis, were employed expressly for research into academic language, a new area of specialization was born, that of English for Academic Purposes (Jordan 1997; Flowerdew & Peacock 2001; Hyland 2006; Bruce 2011; De Chazal 2014). Drawing upon J. Flowerdew (2015c), Basturkmen and Wette (2016:164) define English for Academic Purposes (EAP) as 'the teaching of varieties of English to assist students of all ages to manage the linguistic, conceptual and social demands of academic study, as well as to support the dissemination

and exchange of research and scholarship', thus emphasizing that theory and research are essential compliments to EAP's practical dimension. By including the goal of supporting and sharing scholarship, Basturkmen and Wette tacitly acknowledge the significance of the discourse community, which is a salient feature of EAP research (Bruce 2011).

Also of interest in the above definition is the reference to 'varieties' of academic English since it points to variation as a product of various domains of knowledge and differing language usage. This is especially the case in the higher levels of education. Hyland (2006) supports this view, stating that EAP is applicable from pre-tertiary to postgraduate studies. Because EAP tends to focus on particular skills, such as the structure of written argumentation within a given context (Lee & Subtirelu 2015), it is concerned both with text and content. Thus, it is amenable to large-scale analyses of text in hitherto unexplored disciplines. Though EAP has developed a range of linguistic resources from such research, the fact that EAP courses tend (partly out of necessity) to be taught by subject experts, rather than by trained linguists or writing instructors, is likely to result in the prioritization of content over text. This disadvantages those in need of explicit writing instruction, particularly L2 novices. This being the case, experts could also benefit from an analysis of their

proprietary discourse that offers them a broader linguistic understanding of it.

Because English continues to gain ground as the lingua franca of the academy (Mauranen 2003; Mauranen et al. 2010) and writing is integral to study (Pratt 2011), much EAP research and practice has focused on writing (Hyland 2006; Bruce 2011; Paul & Diani 2015). The result has been an abundance of discoveries about the nature of academic writing, supporting Johns' (2003:207) assertion that 'there definitely *are* conventions in all forms of this writing'. Recalling Halliday and Matthiessen's (2014) notion of context of situation, these 'forms' may be understood as formal types (e.g., textbook, journal article, essay) situated within given disciplines. This indicates that the writing conventions referenced by Johns are likely to exist in unique configurations within each disciplinary text type. To thoroughly investigate the unique configurations within Music Discourse, I chose to analyze two of the most frequent text types produced in Musicology: expert journal articles and novice essays.

EAP research into various types of academic writing has resulted in illuminating analyses of several prominent discourse features: stance (Englebretson 2007; Chang 2012; Aull & Lancaster 2014;

Jiang & Hyland 2015; Crosthwaite & Jiang 2017), discourse markers (Povolná 2012), hedging and boosting (Hyland 1994; Bondi 2015), nominalization (Jalilifar et al. 2017), and use of the passive voice (Tarone et al. 1981), to name a few (Lovejoy 1991; Barton 1993; Carter & McCarthy 2006; Bondi 2008; Parkinson 2013). Already, these findings have contributed to generalizations about the nature of certain disciplinary discourses; for example, that scientific discourse is characterized by technicality, or that philosophic discourse depends more on abstraction (Hyland 2012).

As with ESP, EAP also enjoys a productive relationship with Corpus Linguistics (Paul & Diani 2015). A large swath of recent research from the cooperation of the two is focused on formulaic language in academic writing (Charles et al. 2009; Li et al. 2009; Chen & Baker 2010, 2016; Ädel & Erman 2012; Pérez-Llantada 2014; Alhassan & Wood 2015; Bordet 2015; O'Donnel et al. 2015; Gledhill 2015). Considering the range of investigations into disciplinary discourses that has already been conducted, and the recent interest in formulaic language, the present study applies a corpus-based methodology to identify formulaic items in Music Discourse for the dual purposes of analyzing them and their surrounding propositional content, with the intent of making a novel contribution to the existing body of EAP research.

In this thesis, a focus on Chinese students studying music through English connects the interests of SLA and EAP to compare a specific category of novice with expert writing in a particular area of academic English. From here on, a distinction is drawn between novices and experts, rather than the conventional descriptors of students and teachers. As Solly (2016) notes, considering the challenges for anyone studying an academic subject, the distinction of L1 from L2 is less relevant in EAP than might be the case for other areas of ELT. Instead, the focus on specialized knowledge within a community of practitioners may be better underscored as the distinction of novice from expert. As this study researches Chinese novices, the label L2 will still occasionally be employed when discussing the particular writing challenges these novices face. Reminiscent of Gee's (2014a) recognition, Vickers (2010), following Jacoby and Ochs (1995) and Jacoby and Gonzalez (1991), states that the roles of novice and expert are bestowed upon persons through the ratification or rejection of their contributions to their given academic community: 'The expert-novice relationship is born out of differential experience with and access to community practices' (Vickers 2010:116). In this sense, the roles of novice and expert are recognized through social mediation, thus offering a dynamic scale, imposed by the Community, for the evaluation of a

person's work (ibid.). Writing often serves as the identification badge of expert or novice status.

Because EAP focuses on the language needs of novices, and promotes research to inform expert teaching practice, it often centers on the role of English in academic publication (Basturkmen & Wette 2016). Expert writings may be analyzed to establish criteria for judging, and novice writings may then be compared with such standards to identify what concepts and practices they must learn. Within EAP, the notion of competence covers the expanse of linguistic knowledge necessary for acceptable communication within a given situation (Widdowson 1983; Bruce 2011), which aligns with the aims of SLA (Ellis, R. 2008). Central to this notion is the ability to integrate disciplinary knowledge with the structures of disciplinary writing 'to process and create extended discourse [which] is referred to as *discourse competence'* (Bruce 2008:2). As expert status is recognized through publication (Pérez-Llantada 2014), and publication is a mark of professional achievement, perceptions of discourse competence and professionalism are interwoven (Solly 2016). Thus, publication acts as the merit by which expert status is recognized (Flowerdew, J. 1999a, 1999b; Cho 2004; Jalongo et al. 2014).

Discourse competence can be divided into textual, generic and social dimensions: the first is demonstrated by texts that exhibit grammatical correctness, cohesion and coherence, while the second and third necessitate an understanding of context and community (Bhatia 2004). As discussed above, current volumes dedicated to music writing prize conventional, prescriptive grammar yet ignore cohesion or coherence, not to mention context and community. Consequently, these books address neither fundamental issues of structure nor presentation of disciplinary content. This suggests that recent linguistic insights into the nature of academic writing have yet to infiltrate the Music Department, despite the fact that scholars continue to highlight the necessity of teaching to discipline-specific situations, rather than teaching general academic writing (Coffin & Hewings 2003; Zbikowski 2008; Durrant 2009:165). To redress this situation, Morton advocates teaching a functional approach to language, from the lexicogrammatical to the generic, as a means of inducting students into the culture of a given community (Morton 2010). This suggestion is partially the inspiration for the use of functional analysis in this thesis' methodology. Before this could be successful, however, experts would have to be trained in the basics of a functional approach, which first would require the creation of a functional description of Music Discourse that is accessible for non-linguists. Only in this manner could expert instructors be in a

position to fully assess a novice's discourse competence, which is the threshold of knowledge required to function within the Community (Morgan 2014:37).

2.6 L2 Novices

The lack of research into Music Discourse, my own education as a performer and musicologist, and my thirteen years of experience teaching both English and Music in China combined, prompted me to conduct this analysis of Music Discourse, comparing expert and novice writings both for the benefit of Chinese learners studying music in English (i.e., their L2) and of experts who teach the Discourse. As the world's second most populous nation and second largest economy,[7] China is gaining an ever larger presence on the world stage. Musically, this presence has manifested in a superabundance of piano and other instrumental studies, as students prepare for competitions and examinations.[5] A large number of Chinese students study abroad, approximately half of whom study in English-speaking countries (Ministry of Education of the People's Republic of China 2016).[6] Though not all of these students study music, those who do are certain to impact domestic music education in China and global music education in the near future, which is more than sufficient reason to prepare them for full participation in the international Music community. (N.B.: Specific

statistics on the number of Chinese students studying music in English-speaking educational institutions is not available.)

The above sections have considered the benefits to experts and novices of researching Music Discourse but also the current lack of linguistically informed teaching and learning resources for disciplinary-specific writing instruction. Next, discussion turns to the concept of register and configurations of features within it that can be dependent on discipline. One prominent feature is formulaic language, which is of structural importance in academic writing (Biber et al. 1999; Wray 2008). Thus, it is explored together with the disciplinary content it intersects as a potentially unique configuration, partially constituting Music Discourse as a sub-stratum of the academic register.

2.7 Register

M.A.K. Halliday's Systemic Functional Linguistics, or SFL, models language as nested strata within a socio-cultural context (Butt et al. 2000; Eggins 2004; Crystal 2010; Fontaine et al. 2013; Halliday & Matthiessen 2014; Thompson 2014). Together, these strata form a functional system for the creation of meaning. The result is that 'the process of using language is a *semiotic* process, a process of making meanings by choosing' (Eggins 2004:3). For a writer, choice is

necessitated by the balancing act of reaching personal objectives while simultaneously meeting the expectations of readers (Hoey 2001). This tension is posited as *context of situation:* Malinowski's (1923) functionalist conception of language as *situated in and by its use* (Campbell 2006) [emphasis added]. *Context of situation* was taken up later by Firth (1969), who used it to develop his thoughts on collocation, as well as his idea that language is composed of constituent levels (Léon 2006). From his studies with Firth, Halliday developed his own graduated model into the theory of Systemic Functional Linguistics (Cummings 2006). 'When texts share the same context of situation to a greater or lesser extent, they will share the same experiential, interpersonal and textual meanings and so they belong to the same register' (Butt et al. 2000:9).

In SFL, register is understood as a variety of language associated with a given situation in which it is employed to achieve specific goals (ibid.; Grimshaw 2003; Halliday & Matthiessen 2014; Thompson 2014). 'The description of register covers three major components: the situational context, the linguistic features, and the functional relationships between the first two components', such that the features are understood as functional when filtered through context (Biber & Conrad 2009:6). As 'clusters of associated variants', registers are the means of communicating particular

content to a particular discourse community; therefore, a register consists of features crucial to a chosen construal of its particular content (Halliday 2001:158). From the perspective of SFL, a novice's development of overall linguistic ability and particularly of discourse competence is positioned within the acquisition and development of register (Yasuda 2011), and therefore its features, which is why L2 novices ought to be introduced to register as a central characteristic of language (Biber & Conrad 2002).

The scope, depth and sheer volume of academic endeavors have generated a large and particular cluster of variants, rendering 'present-day professional academic writing…one of the most distinctive registers in English' (Biber & Gray 2010:18). Because Music Discourse is a sub-stratum of this register, it was necessary to delimit its context of situation for the purpose of identifying its distinctive linguistic features before selecting representative texts with which to analyze it. The characteristics of this context of situation are extralinguistic and include:

the participants, their relationships, and their attitudes toward the communication; the setting, including factors such as the extent to which time and place are shared by the participants, and the level of formality; the channel of

communication; the production and processing circumstances (e.g., amount of time available); the purpose of the communication; and the topic or subject matter (Biber & Conrad 2015:175).

Working through this list in reverse, Music Discourse within the academic register may be explicated as follows: the shared purpose of the Discourse is communication and dissemination of music analysis and description to the Community through peer-reviewed publications and even some assessed student texts; Discourse texts tend toward high degrees of formality and lexical technicality; texts are mostly shared among a relatively small numbers of experts, whose degrees and academic positions rank them as colleagues, and who are spread across the globe yet connected by shared interests; English is the academic language of the Discourse. Among these, the characteristics distinguishing this sub-stratum from the larger academic register are its disciplinary content, including terminology. Thus, configurations of features unique to the Discourse are to be sought among these features. These configurations are first investigated as recurrent sequences of words. Thus, the present analysis focuses on formulaic language.

2.8 Formulaicity

When analyzing discourse, N.C. Ellis (2003) agrees with Sinclair's advice to depart from the *idiom principle*, the notion that texts tend to be constructed from certain formulae as opposed to wholly original choices (Sinclair 1991:110). According to this principle, text can be analyzed to discern which patterns are typical of a given type and how such patterns can be generalized to register, or in the present study, a registerial sub-stratum. Following this idea, the present analysis of Music Discourse departs from a study of formulaicity. As a sub-stratum of academic writing, Music Discourse inherently shares certain configurations of features with the academic register as a whole. Nonetheless, the disciplinary content of this Discourse introduces the means of altering those configurations or creating new ones unique to Music Discourse. Thus, identification and description of these patterns requires analysis of pervasive textual features. Formulaic items appear in every language variety but are found to be especially common in the academic register, which suggests a functionality within academic text (Biber & Conrad 2009).

As Halliday and Matthiessen (2014:32-36) note, usage peculiar to a subject becomes discernible only at a sufficient degree of delicacy, that is, specificity. Considering this, a detailed inspection of

formulaic language within Music Discourse has the potential to identify patterns of usage specific to this Discourse. As such study requires a bulk of authentic texts, a corpus-based methodology is the most efficient means of extracting these patterns, since corpora can be constructed from texts in a given register to mine for frequent items within that register. Thus, corpus offers a powerful tool for comprehensive description of certain textual features. For instance, 'Corpus linguistics demonstrates that much of communication makes use of formulaic sequences' (O'Donnel et al. 2015:83-84). Extracting recurrent structures or words from a corpus is a simple matter of identifying them by frequency. This ability to illuminate frequency means that corpus offers the preeminent means of identifying linguistic items relevant to teaching and learning (Ellis, N.C. 2014).

2.9 Corpus Linguistics

Corpus Linguistics (CL) 'is an area which focuses upon a set of procedures, or methods, for studying language' (McEnery & Hardie 2012:1). Descriptions such as this view Corpus as primarily methodological in nature; however, this overlooks the fact that Corpus offers access to specific types of linguistic data (e.g., formulaic items), which in turn facilitates inductive theorizing about the nature of language (Halliday & Matthiessen 2014:53). It is

perhaps on account of such competing claims that Corpus Linguists are not in universal agreement as to the precise definition of their field; indeed, an explanation that encompasses all activity deemed part of CL is difficult to formulate. Nonetheless, the most basic definitions would note that a corpus is a compilation of machine-readable texts, selected according to specific criteria, for the purpose of addressing particular research questions (McEnery & Hardie 2012), or in Conrad's (2002:76) concise formulation: 'A corpus is a large, principled collection of naturally-occurring texts that is stored in electronic form (accessible on computer)'.

The most significant contribution of Corpus to linguistic research is its ability to process enormous amounts of naturally occurring, or non-invented text. This makes it a valuable tool for SLA, one tenet of which holds that research should be conducted using authentic data (Gilquin 2015). As a result of this potential, corpora are built for an array of practical ends, such as investigation into varieties of academic discourse or comparison of expert and novice language usage. Of course, any such analysis ought to be principled; thus, a corpus should only be assembled having first clearly defined the research questions it is expected to address (McEnery 2004). To ensure applicability to a given research problem, corpora are constructed as a carefully chosen set of texts. When designed

according to principled criteria, corpora boast several advantages. As a tool for empirical research, Corpus Linguistics affords a reduction in researcher bias (Baker 2006). Its reliance on computers renders it more efficient and accurate than human computation, and it offers a convenient means of replicating research (McEnery et al. 2006). Teubert (2004) further argues that CL is uniquely positioned to give insights into meaning because it deals in natural language, which is an outgrowth of shared experience, though the veracity of this claim diminishes in inverse proportion to the amount of manual calculation and qualitative analysis required by a given set of research questions. CL is also particularly effective for the identification of linguistic items sufficiently frequent as to warrant learners' attention (Ellis, N.C. 2014). The present corpus-based[10] study harnesses the advantages of CL's quantitative methods to conduct both descriptive and explanatory analyses (Thornbury 2010).

Given its assets, a cautionary reminder is needed regarding the limitations of Corpus Linguistics. Though CL is able to reveal various aspects of language, it is not in possession of an independent explanation thereof (McEnery et al. 2006). Thus, it is only employed methodologically in this study. Because corpus is empirical and inductive, it must hedge its claims about any language system with

the caveat that it may speak confidently only of that data which it has directly investigated; consequently, CL offers a view of language behavior under particular conditions, which in turn permits an opportunity for inference of a larger reality (Stubbs 2007a). These limitations provide ample reason to use corpora in conjunction with other analytical methods, as is done in this study.

As corpora are often purpose-built, they come in several varieties. General corpora, which are usually large, cover a range of genres and knowledge domains in order to offer the broadest possible view of a given language (Biber & Conrad 2002; Biber et al. 1996; Cheng 2012; Cortes et al. 2015; Crawford & Csomay 2016; Kennedy 1998). By contrast, specialized corpora are typically smaller and are constructed to represent a sub-language, focused either on domain or genre (Flowerdew, L. 2004; Koester 2010). They are therefore well equipped to facilitate discourse analysis by offering a direct view of a given epistemological perspective using a relatively manageable amount of data (McEnery et al. 2006). One type of specialized corpora, a learner corpus, focuses on language acquisition, either longitudinally or in cross section for the benefit of L2 learners (Gilquin 2015; Granger et al. 2015; McEnery et al. 2006). Such corpora can offer instructors an empirical basis for

language teaching by illuminating aspects of authentic language that require attention (Granger 2009).

While the construction of learner corpora is a relatively recent phenomenon (Granger 2009; Granger et al. 2015), already several prominent examples have made valuable contributions to corpus research, including the Longman Learner Corpus (LLC), Cambridge University Press (CUP) Learner Corpus, and the International Corpus of Learner English (ICLE). These and other similar corpora have assisted the determination of which language features to teach, such as formulaic sequences (Nesselhauf 2004). Still, the benefits of this research have yet to affect some teaching materials, as is the case with music writing. In fact, it appears at present that even music terminology dictionaries remain insulated from corpus research. This is an area in need of further study for the benefit of novices, particularly L2 novices, since music terminology presents the learner with a dizzying array of terminology, much of which is not of English origin, being drawn from Italian, French, German, Latin, Spanish and a handful of other European languages.

As noted above, specialized corpora are used to research specific genres or domains of knowledge. Because corpora can be tailored to specific research concerns, specialized corpora are inherently well-

suited to discourse analysis (Biber, Connor & Upton 2007). '[C]orpus-based methodologies lend themselves well to answering the questions relevant to disciplinary specificity' (Friginal and Hardy 2014:26). Meunier (2002) illustrates this benefit of CL by noting how corpus studies have revealed the lexical, grammatical, syntactical and discoursal fingerprints among various forms of writing. To analyze Music Discourse thus, two specialized corpora were built for this study: one of expert writings in the form of academic journal articles, the other a learner corpus of novice essays. This provides a quantitative approach to the identification of recurrent features across texts (Flowerdew, L. 2014; Conrad 2002) to highlight the complexity of these seemingly simple structures (Altenberg & Granger 2001).

A final caveat is given here on the limits of learner corpora. Because they are typically much smaller than general ones, learner corpora produce data that must be analyzed cautiously. It is necessary to appropriately hedge claims made on such a basis since small collections of text offer a limited view of language use. Thus, the only claims that can be made confidently are those for the data in the corpus. In other words, it is difficult to extrapolate tendencies from such limited data, though this may be somewhat more possible given a qualitative approach that analyzes the data in context. Still,

the smaller size of such corpora may impose limitations to representativeness and balance, as discussed below.

2.10 Formulaic Language

Until this point, the designation *formulaic* has been used to reference the repetition of sequential linguistic items. Such sequences may be any series of contiguous words that exist as a unit and which seem to be stored and retrieved from memory as such (Wood 2015). As mentioned, the expert authors of music writing guides advocate for grammatical accuracy and clarity of purpose in disciplinary writing, yet the resources they offer make no mention of the structures that bind sentences together as whole texts. This is unsurprising, however, as it simply confirms that disciplinary experts lack specialized linguistic knowledge. As a consequence, they do not (and likely are unable to) advise novices how to work from grammar to discourse in their writing. Because formulaic sequences are ubiquitous in academic writing, an understanding of these sequences, their content and function, could benefit both novice and expert writers.

Study of formulaic language can be traced to Firth's (1957) work on collocation and his influence on subsequent scholarship (Léon 2006). For instance, Sinclair's *idiom principle* — the notion of

partially fixed phrases that function as single units within text —
looks back to Firth's work, while also often serving as a point of
departure for later definitions of formulaic language (Sinclair
1991:110). The recurrent nature of such sequences has led to a
general acceptance that 'There is undoubtedly some sort of
relationship between frequency and formulaicity, both in the sense
that *some* formulaic sequences are very frequent, and that
formulaic output is frequently called upon' (Wray & Perkins
2000:6-7). The fact of this recall may be of particular interest to
novices who depend on reuse of certain memorized formulations to
generate text.

By now, the sheer volume of research into formulaic language has
generated a detailed account of its characteristics. Formulae exist
across registers and domains; they generally are: '1. Multi word, 2.
Have a single meaning or function 3. [and are] prefabricated or
stored and retrieved mentally as if a single word' (Wood 2015:3).
Indeed, so ubiquitous are these formulae that they have come to be
regarded simply as multi-word sequences (Wray 2008). The facts
that formulae are pervasive, 'are not amenable to lexical and
structural re-formulations', and 'tend to occur in particular styles of
language tied to particular communicative situations' (Corrigan et
al. 2009:XIV), suggest that these sequences can be viewed as an

indicator of language fluency achievement. In turn, the realization of their ubiquity has been cited as a reason to study both their function and context (Mollet et al. 2011; Qin 2014; Ebeling & Hasselgård 2015; Hatami 2015; Liu & Nelson 2016). Some of this research has shown that such sequences and their configurations may be partially differentiated according to context and discipline, again indicating that these formulae are indispensable to the achievement of discourse competence (Wray 2012). SLA research supports these conclusions, having demonstrated that learners acquire language over time through a continual process of aggregating smaller units into larger ones (Ellis, R. 2015). This puzzle-completion approach is aided by the input of these word sequences, as such formulae provide examples of the target language's grammatical structure and experiential content, assisting learners to link pieces of semantic content according to grammatical norms (Myles 2012). Though a top-down view of language is needed to more fully comprehend a discourse, novices seldom survey language from this vantage. Consequently, it is understandable that they should seek to memorize ever larger and more complex strings of words, thereby building upward.

The significance of formulaicity in academic writing has been demonstrated by numerous studies (e.g., Hyland 2008; Staples et

al. 2013; Kashiha & Heng 2014; Pérez-Llantada 2014; Alhassan & Wood 2015; Peters & Pauwels 2015), many of which have produced lists of the formulae found in such writing (e.g., Biber et al. 1999; Biber et al. 2004; Simpson-Vlach & Ellis 2010). As Pérez-Llantada (2014:92) observes, 'formulaicity is a key feature of the academic written register across language variables', at least in the sense that it is frequent. Because of their centrality to academic writing, Wray (2008) asserts that formulaic sequences promote survival within an academic community, an argument echoed in Hatami's (2015:116) description of formulae as 'zones of safety' for learners. To some extent, this is undoubtedly the case, though as that reliance deepens and expands, it is increasingly likely that novices would begin to overuse the formulae most familiar to them. This does not negate the feeling of safety they may offer, but it does diminish their effectiveness for assisting novices to improve their writing. This potential danger notwithstanding, Wood reiterates the importance of formulae to academic discourse, arguing that 'Formulaic sequences are, in essence, a major part of the foundation of successful academic writing skills because they comprise the basic elements of academic discourse and are specific to particular disciplines, registers, and genres' (Wood 2015:103). This argument provides a rationale for researching formulae as some are particular to specific disciplinary discourses.

As the understanding of formulae in academic writing deepens, it has become increasingly apparent that these are indispensable to the achievement of competence in disciplinary writing (AlHassan & Wood 2015). This in turn has prompted awareness of their pedagogical value. Because formulaic sequences present a significant challenge to novices, they have become of interest to EAP researchers (Hiltunen & Mäkinen 2014). For example, the above study (ibid.) investigated differences between novice and expert writings in the domain of business and economics, using the Academic Formulas List (AFL) of Simpson-Vlach and Ellis (2010). Having mined their research corpus for formulaic sequences, they analyzed individual student writing samples to identify examples of novice usage (Hiltunen & Mäkinen 2014). Formulaic sequences are so common in L2 production that they may even permit novices to appear more competent than they actually are (Myles 2012). Comparison of usage frequencies between expert and novice production can suggest areas to be addressed in teaching practice. This type of comparison can also yield valuable insights into the structure of Music Discourse by revealing the patterns found in expert writing as an exemplar for novices. Of course, given the sheer number of possible formulaic sequences, those most typical of a given discipline should be carefully prioritized for research and

teaching lest the learning burden overwhelm the novice (Hatami 2015).

2.11 Lexical Bundles

The preceding discussion has expounded formulaic language by defining it, noting its frequency and pervasiveness in academic writing, and by discussing its potential value for disciplinary-specific teaching. The corpora in this study were mined specifically for one type of formulaic sequence, lexical bundles. These recurrent features are the largest items extractable from a corpus, and thus they offer the broadest possible quantitative view of formulae the corpus can provide.

Chen and Baker (2010:30) have catalogued the variety of terms used to designate such sequences of co-occurring words: 'clusters' (Hyland 2008; Schmitt et al. 2004; Scott 2017, *WordSmith Tools*), 'recurrent word combinations' (Altenberg 1998; De Cock 1998), 'phrasicon' (De Cock et al. 1998), 'n-grams' (Stubbs 2007c; Anthony 2018, *AntConc*), and 'lexical bundles' (Biber & Barbieri 2007; Cortes 2002)'. In this study, I employ the term 'lexical bundles' on account of its connection to the work of Biber and his colleagues, whose models of this phenomenon inform the present study, and also for the term's currency in much of the literature (Chen & Baker 2010).

Lexical bundles (LB) represent a particular form of formulaic language: a contiguous sequence of words that is frequent and has either a single meaning or function (Biber et al. 1999). Altenberg (1998) originated the investigation of these bundles, concentrating both on their frequency and function. They occur across disciplines, with specific intersections of relative frequency and type serving as markers or characteristics of a particular discipline (Wood 2015). As indicators of specific disciplinary discourses, lexical bundles have gained prominence in language learning research (O'Donnel et al. 2015).

Comprehensive research into the form and nature of LB was published in *The Longman Grammar of Spoken and Written English*, which defines lexical bundles as 'recurrent expressions, regardless of their idiomaticity, and regardless of their structural status' (Biber et al. 1999:990). This last point highlights the fact that these bundles tend to straddle structural boundaries, as well as functional ones. Despite this, Biber et al. (2004) state that LB offer a particularly useful view of the structural and semantic content of a given register, such as academic discourse, because they 'have been shown to be discipline-bound, with each discipline or academic community having its own unique recurrent word-combinations' (Qin 2014:230). Wood (2015:165) elaborates on this: 'The remarkable

and paradigm-shifting effects of the discovery of lexical bundles have uncovered the internal working of academic discourse, providing us with an observable and tangible element of language which is woven deeply into the fabric of discourse'.

While this research has found that bundles are bound to particular disciplines and thus offer insights into a given register, the extent to which those insights are structural is debatable. Indeed, Wood's claim that these bundles lay bare the internal mechanics of the academic register is overstated. Though some bundles provide formulae relevant to stance, discourse structure or reference (Biber et al. 2004), the fact that they often straddle grammatical boundaries tends to obscure the structures of discourse, perhaps even to the same extent that these bundles reveal its formulaic features. To take full advantage of the lexical bundle perspective of discourse, it is necessary to view LB in the context of their surrounding content. Only in this way can bundles be properly understood as patterns of recurrent sequences within a larger grammatical context. Such a binocular view requires the application of qualitative manual analysis to the quantitative results of the corpus, yielding a powerful lens through which to observe the intersection of formulaicity and content in a discourse.

Not only does this binocular perspective offer a view of lexical bundles in their grammatical context, but it also opens a window onto discourse content, particularly as many bundles tend to be 'referential' in nature; that is, connecting to, or joining propositional content (Biber et al. 2004). Thus, the meaning of a bundle in a given discourse typically depends on its context, in much the same way that words rely on context for meaning. For present purposes, I refer to the content on either side of a bundle as *ambient content* filling a 'slot' (a term employed by the *Longman Grammar* 1999). Observing bundles with their surrounding content facilitates a functional analysis of the interaction between formulaicity and originality, which is useful for producing a detailed map of disciplinary-specific patterns of content. By observing bundles in context across the corpus and analyzing their interaction with ambient content from a systemic functional perspective, a map can be produced of the textural patterns constituting Music Discourse. Thus, I have followed Halliday and Matthiessen, who advocate the application of theory to corpus data to balance macro and micro views of language:

> We would argue for a dialectical complementarity between theory and data: complementarity because some phenomena show up best if illuminated by a general theory

(i.e. from the 'system' end), others if treated as patterns within the data (i.e. from the 'instance' end); dialectical because each perspective interpenetrates with and constantly redefines the other (Halliday & Matthiessen 2014:53).

In the present study, the approach from instance is constituted by the lexical bundles mined from the corpora with their ambient content, while the approach from system consists of the analysis of types of bundles and the patterns they form with content.

Because they seldom form functionally complete grammatical units, lexical bundles are not typically mentioned in grammars, the notable exception being *The Longman Grammar* (Biber et al. 1999). Therein, bundles are classified according to their constituent word classes, demonstrating that many of these bundles are composed largely or entirely using closed word classes, the inherent limits of which readily explain their frequency. In contrast to this form of classification, which offers little insight into the grammatical function of lexical bundles, a systemic functional analysis reveals that bundles consist either of multiple complete or fragmented groups (e.g., nominal and verbal groups), or less often, a single grammatical unit, such as a complex preposition (see Halliday &

Matthiessen 2014:423). In functional grammar, a group is 'a combination of words built up on the basis of a particular logical relation', such as the words forming a nominal group (ibid.:362).

Whereas bundles consisting of a single, intact group are easily analyzed, those fragmented in various ways are not. Among the fragmentary bundles, those consisting of a fragment of a single group may have more than one grammatical function in differing contexts. Such is the case, for instance, with bundles that initiate prepositional phrases — though it shares the same rank as a group, 'a phrase is a contraction of a clause' (Halliday & Matthiessen 2014:9, 362-363), which may serve as clausal adjunct or as Postmodifier in a group (ibid.:424). Bundles constituted of fragments of multiple groups operate differently, exposing the syntagmatic joinery, as it were, of grammatical units of equal or unequal rank.

2.12 Lexical Bundles in Novice Texts

Some studies have questioned previously accepted wisdom that generic skills, such as forms of argumentation, are transferable across disciplines; instead, it is increasingly suggested that such skills can only be effectively developed within a proper disciplinary context (Swales 1990; Gimenez 2011). As already noted, however,

this proves challenging because expert instructors usually do not possess sufficient linguistic knowledge of their own discourse. Considering this, provision of the most common formulae, such as lexical bundles, along with their ambient disciplinary content, could offer a means of support for both teaching and learning. In support of this claim, Crossley et al. (2012:215) argue that knowledge of bundles contributes specifically to discourse competence: 'One reason for the centrality of [lexical bundle] knowledge as a mark of language acquisition is the notion that [bundles] contain both the paradigmatic and syntagmatic features' of lexical and syntactic knowledge, respectively. Furthermore, research into formulaic language suggests that lexical bundles may be advantageous instructional aids if they are taught with the content words that often surround them (see Nesselhauf 2003). In this manner, novices may learn how to competently structure discipline-specific meanings in their texts. Learning to deploy lexical bundles successfully, however, presents challenges for all novice writers. Novices sometimes write the same bundles as professionals, yet employ them in unusual or inappropriate ways (Wray & Perkins 2000). This challenge looms larger for L2 novices (Yuldashev et al. 2013), which has prompted research into the value of teaching lexical bundles to L2 novices (Ellis, N.C. 2012, 2014). It may be that novices' interest in disciplinary content leads them to focus on the acquisition of

formulae, such as LB. Should it be the case that novices focus solely on acquisition, though, they might yet find themselves unable to deploy bundles competently.

Of course, simply generating lists of lexical bundles, even organized taxonomically, would prove insufficient to help either experts or novices, since they would not possess an understanding of their use, nor especially of their interactions with content throughout a text. Moreover, this would constitute far too much information for memorization. Therefore, research into applicability is much needed, as 'formula in context is what is pedagogically relevant' (Simpson-Vlach & Ellis 2010:502). Hence, for Music Discourse, it is necessary to investigate how LB interweave with propositional content to construct knowledge. Hunston and Francis (2000) surmise that formulaic items form patterns in texts, a conjecture that is supported by the present research. Perhaps the intersections of lexical bundles and content provide a vital and heretofore unrecognized link between the teaching of grammar and discourse for certain groups of novices. This idea is supported by AlHassan and Wood (2015:52): 'The view of academic discourse as replete with formulaic sequences implies that academic writing skills surpass the mastery of lexicon and syntax to encompass the

successful implementation of these sequences that are viewed as the building blocks of academic discourse'.

The above argument is further, albeit indirectly, defended by Altenberg and Granger's (2001) contention that much information pertinent to language teaching can still be gleaned by employing simple corpus tools (e.g., lemmatizers, concordances and formulae) for the purpose of revealing students' strengths and weaknesses. One obvious novice weakness is the uncertainty of how to paraphrase. Since failure to do so may result in plagiarism, novices could greatly benefit from explicit instruction about which recurrent strings of words are replete in a given discourse. This could assist them in the identification of non-formulaic language, which is more likely to be taken verbatim from a source.

2.13 Taxonomy of LB Discourse Function

Even though corpus tools have gained a certain popularity among Discourse Analysts for their ability to efficiently sort through an enormous amount of text (Flowerdew, J. 2016), offering a broad overview of lexical phenomena across whole corpora, the sheer volume of data such corpora are able to produce presents a significant impediment to any qualitative research conducted with that data. Aside from the daunting task of manually analyzing all

data, it is necessary to collate findings to reach a global vantage of those findings. This calls for the creation of some type of classification or taxonomy. This is particularly necessary for the present research because of the inclusion of 3-word lexical bundles. Despite the prevalence of studies focused on 4-word bundles in writing (Chen & Baker 2010), 3-word bundles were also included herein because their pervasiveness suggests that they are likely to include more disciplinary terminology than longer bundles. *The Longman Grammar* (1999:993) lists other such characteristics of LB. Among these, several are pertinent to this analysis of Music Discourse: 3-word bundles are more numerous than 4-word ones by a factor of 10; bundles in academic writing typically consist of noun groups and prepositional phrases, ending in function words, and therefore tend to be nominal rather than clausal (ibid.:1,000); when bundles cross structural boundaries, such as clauses, the following slot often contains the content specific to a given situation (ibid.:995, 999); bundles typically overlap each other, forming longer bundles (ibid.:999). Several of these characteristics support the need for the classification of lexical bundle data, such as their tendency to overlap one another and their grammatical similarities. Furthermore, the presence of noun groups, especially in 3-word bundles, demonstrates a link to content that adds to the need for a classificatory system.

Though it can be problematic to create a taxonomy for lexical bundles, Simpson-Vlach and Ellis maintain that it is crucial for pedagogical purposes, urging that bundles be studied in context, as doing so sometimes expands a given taxonomic category (Simpson-Vlach & Ellis 2010). Howarth (1998) put forward one of the earlier taxonomies for phraseology, but this has been replaced by more recent work, particularly that of Biber et al. (1999) in the following year. While *The Longman Grammar* grouped lexical bundles in academic prose into 12 structural categories (Biber et al. 1999:1014-1015), a later study by Biber et al. offers three structural types for bundles, incorporating fragments of verb groups (e.g., 'this is a'), relative clauses (e.g., 'that there is'), noun groups or prepositional phrases (e.g., 'one of the', 'at the end of') (2004:384). These structural types were then used to propose a provisional tripartite taxonomy by function.

The resulting three functions proposed for lexical bundles are stance expressions, discourse organizers and referential bundles, defined as follows:

> Stance bundles provide a frame for the interpretation of
> the following proposition, conveying two major kinds of

meaning: epistemic and attitude/modality [either personal or impersonal] (Biber et al. 2004: 389).

Discourse organizing bundles serve two major functions: topic introduction/focus and topic elaboration/clarification (ibid.:391).

Referential bundles generally identify an entity or single out some particular attribute of an entity as especially important, [and are common in academic prose] (ibid.:393).

Furthermore, there is a consistent correlation between structural and functional dimensions in frequent bundles that alludes to an interweaving of structure, function, register and form in the construction of a given discourse; thus, evidence continues to mount suggesting that lexical bundles are 'a basic linguistic construct', heavily dependent upon context of situation (ibid.:398).

Though alternate taxonomies have been proposed — Hyland (2008a:49) predicates his on Halliday's metafunctions, while Cunningham (2017:76) coins his own tripartite categorization:

aboutness, coherence, variable level discourse — Biber et al.'s (2004) is adopted in the present research on account of its clear tripartite division, its delineation of various referential expressions in academic discourse, and because it is established among researchers (Wood 2015; Wray 2008, 2012).

Lexical bundles 'can be regarded as structural 'frames' succeeded by a 'slot'. The frame functions as a kind of discourse anchor for the 'new' information in the slot, telling the reader how to interpret that information' (Biber et al. 2004:399). Such slots are likely to contain the experiential content necessary to reinterpret bundles within a disciplinary context. Given the fact that bundles can cross structural boundaries, it is reasonable to consider the information preceding such bundles as another slot.

Because this study investigates the intersection of lexical bundles with disciplinary content in Music Discourse, I have narrowed the analytical focus to referential bundles, since of the three categories in Biber et al.'s taxonomy, this one is most likely to introduce nominal groups, which in turn contain the content, or Things, of the Discourse.

2.14 Discourse Analysis

Corpus and SFL can be employed complementarily to conduct discourse analysis. The former provides a powerful data-processing tool with which to sort vast amounts of authentic data, offering a theoretically 'thin' description, while the latter adds a complex view of language that can be used to create a theoretically 'thick' description (Thompson & Hunston 2006:2), a descriptive quality it shares with EAP (Hyland 2006). Given this complementarity, I have chosen to triangulate quantitative (Corpus) and qualitative (SFL) approaches so as to generate both breadth and depth of insight into the nature of Music Discourse. Notwithstanding this triangulation, the fact remains that a given analysis 'cannot presume to have exhausted the meaning of the discourse' (Martin 2012/2009:357). Consequently, certain aspects of Music Discourse have been omitted from this analysis, most notably its multimodal dimension: the inclusion of score excerpts.

The term *discourse analysis* originated in 1952 with Zellig Harris ('Discourse Analysis: A Sample Text', as given in Todd 2016). Broadly speaking, discourse analysis is the investigation of 'language-in-use' (Brown & Yule 2000). Adding texture to the description of discourse as action (Lillis & McKinney 2003), Fetzer (2012:454) draws attention to its structural relationships, stating:

'Discourse is fundamentally concerned with the nature of the connectedness between parts and wholes'. Presently, discourse analysis (DA) blankets a range of methodological approaches, largely separable into three categories: the study of language use, the study of linguistic structure larger than a sentence, and the study of social practices (Tannen et al. 2015). The second of these, text analysis, is occupied with the investigation of relationships within a text (Sanders & Sanders 2006), such as co-reference across spans of text (Grimshaw 2003). Other relationships of interest for text analysis include context and content, as well as cohesion and coherence (Graesser et al. 2003:2). Regarding the former, content analysis is one of the oldest established methods of text analysis. Subsequently, it has grown difficult to define because it has roamed across expansive categories and concerns. As a result, content analysis has outgrown any single definition to encompass multiple research strategies rather than specific methodologies (Titscher et al. 2000). The present study investigates content through configurations of lexical bundles and ambient propositional content.

The preceding analytical concerns address questions germane to the study of both expert and novice texts. Of special concern for the latter are those related to competence. As discussed previously,

competence is judged on a continuum from novice to expert, often in response to perceptions of typicality. For instance, novices may have retained a store of lexical bundles, yet still be unable to deploy them in ways typical of, and acceptable to, a given disciplinary community (Barton 1993). Ädel and Erman (2012) reiterate the point that competence may be achieved by learning the conventional deployment of formulae specific to a particular disciplinary register. Various studies have addressed this question. For example, investigations comparing L1 and L2 novices have found that both groups' overall knowledge of bundles tends to be somewhat restricted to discourse markers, to the neglect of referential bundles (Nekrasova 2009). This clearly poses a difficulty, as referential bundles are known to be common in academic writing (Biber et al. 1999). Hyland (2008) also stresses the need for additional exploration of how bundles function differently from expert to novice writings, which again touches upon the question of typicality. In fact, Chen and Baker's (2010) research demonstrated significant differences in the use of lexical bundles between academics and students; this finding suggests that a similar situation may be found between expert and novice Music scholars. Finally, Wray (2012:244) notes that research into formulaic language has 'barely touched on the question of communication — the meanings we want to make, and the subtle judgments we

need to engage in, so as to make them effectively'. All of this intensifies Bestgen's appeal to replicate research into formulaic language on various corpora to continue yielding vital information for the teaching of writing (Bestgen 2017). Thus, high-frequency items identified within the register of Music Discourse are analyzed to see how they interact with disciplinary content to construct text (Thomas 2015).

These relationships present a complex textual web for investigation. To address this situation, Fetzer (2012) advises conducting analysis from both the macro and micro levels to simultaneously accommodate quantitative and qualitative dimensions. This accords with Martin's (2012/2009:335) advice to 'shunt around' between levels of language when analyzing text, though a linguistically informed rationale should be imposed in order to generate a disciplined approach. In this study, formulaic items represent the micro level, while the analysis of how those items interact with content to form the Discourse represents the macro. Herein, the quantitative dimension, enabled by corpora, represents the first stage of analysis. This phase cannot stand on its own, however, as LB do not contribute substantially to an understanding of an overall discourse (O'Donnel et al. 2015). Still, the use of corpora to conduct a discourse analysis accords with Biber et al.'s (2007:2) statement

that 'Corpus linguistic studies are generally considered to be a type of discourse analysis because they describe the use of linguistic forms in context', provided such studies range beyond the extraction of data from the corpus, as advocated by Halliday and Matthiessen (2014:53). Because such quantitative corpus studies cannot offer a detailed view of how corpus data interacts within texts to build discourse, a second stage of qualitative analysis is needed to reveal the content of the Discourse.

This type of multi-staged discourse analysis, ultimately oriented towards patterns of discourse, aims to provide a map that can aid the development of discourse competence through a more profound understanding of text construction. Lexical bundles contribute to textual coherence by bridging structural units (Hyland 2008). Since bundles offer a short-cut to realizing communicative goals, they can be a valuable resource for developing novices' discourse competence (Wray 1999).

Considering the above, the following research questions were formulated in order to produce a map of Music Discourse that offers a broad view of its patterns and content, which ultimately reveal its epistemological concerns.

2.15 Research Questions

1. Which general and disciplinary-specific lexical bundles are constitutive of Music Discourse as a sub-register of academic discourse?

2. What ambient propositional content is connected to these lexical bundles?

3. What epistemological concerns of Music Discourse are instantiated through the intersection of lexical bundles with this propositional content?

4. How competently do novices deploy these bundles and do these writers instantiate propositional content similar to that of expert writers?

3 Methodology

3.1 Introduction

Because this study focuses on lexical bundles and their ambient content, a substantial portion of both Corpora had to be analyzed to highlight the propositional content in the surrounding slots. The sheer bulk of text to be examined necessitated a methodology that is at once detailed and comprehensive. This resulted in numerous steps, which combined formed a complex process of sorting, collating and analyzing data.

To accomplish this, I employed a mixed methodology. The quantitative axis is provided by structural and functional analyses of two bespoke Corpora and the qualitative by a cartographic analysis of ambient content. The qualitative dimension is then enhanced by analysis of discoursal patterns of formulaicity and content as they operate across complete texts. This approach combines the empirical nature of Corpus Linguistics with the descriptive nature of discourse analysis. Thus, this study adopts as its starting point Biber et al.'s (2004) exploratory approach of mining corpora for frequent lexical bundles. Choosing lexical bundles as a point of departure was a means of trawling with the largest net available to gather as much data as possible. Given the amount of data to be processed, the mixed approach to analysis of that data, and the desire to map

Music Discourse from a novel perspective — namely, the intersection of lexical bundles and content — it was unavoidable that any methodology employed should be complicated. Thus, the stages of the methodology are outlined below.

Before building the two Corpora herein, a pilot study was conducted with a sampling of articles from reputable journals. This informal first step immediately confirmed that lexical bundles in such texts include disciplinary-specific terms, a finding that supported the decision to undertake the present study.

As this research is corpus-based, it was first necessary to choose an appropriate corpus software, and then to weigh considerations of text types for experts and novices, representativeness of these texts, sampling frame and balance, as well as frequency and dispersion before constructing the two Corpora (Section 3.2). Data was then collected, consisting of expert journal articles and novice essays, the latter of which proved challenging to procure (Section 3.3). The texts gathered were prepared for inclusion in the Corpora. Lexical bundles were extracted from both Corpora and the resulting concordance lines copied into word processing documents and spreadsheets for ease of analysis. Lines were analyzed individually according to a rubric, then collated in two spreadsheets, one per

Corpus (Section 3.4). This facilitated large-scale analysis of grammatical structure and function, bundle and content interaction across texts, comparison of expert and novice usage, and an overview of the discipline's epistemological concerns (Section 3.5).

3.2 Corpora

3.2.1 Representativeness

'The beauty of corpora is that they can be created with a purpose' (Friginal & Hardy 2014:30). Building corpora of any variety entails consideration of what will be included and why. The three main concerns guiding these decisions — representativeness, sampling and balance — are discussed below in relation to the Corpora in this study.

For a corpus of writing, the range of variables related to representativeness renders this a difficult concern to fully address (Nelson 2010). Since no corpus can claim to provide an exhaustive account of a language, texts that reflect the communicative functions of a given discourse should be prioritized to achieve maximum representativeness (Sinclair 2005). Thus, it was necessary to consider the texts produced and consumed by the Musicology Community, as well as their relative distribution. As rigorous selection methods are integral to the design of corpora

(Graesser et al. 2003), Sinclair offers the following general criteria: 1) choose and apply structural criteria to the corpus framework; 2) decide which text types are available and should be included, then prioritize them; 3) estimate the size, quantity and importance of the texts to be gathered (Sinclair 2005:7). Biber (1993) agrees and further advises that structural criteria be chosen first by register, after which criteria can be decided according to text types and their distribution, as well as the number of texts and tokens that will be included in the corpus. Because '[t]he register perspective characterises the typical linguistic features of text varieties, and connects those features functionally to the situation context of the variety' (Biber 2010:242), this approach is advantageous for constructing corpora sufficiently disciplined both to support claims made based on its findings and to strengthen the validity of those claims extrapolated to the entire register under investigation. In keeping with this criteria, I selected texts from the academic register, focusing specifically on the discipline of Musicology.

The specialized Corpora in this study were built from the contexts of Musicological writings from expert journals and novice essays. Such a contextual demarcation permits extensive research of, and insight into, the context of the items in the corpus (Flowerdew, L. 2004). Moreover, the Corpora were constructed according to a narrow focus

on text type (articles and essays), subject (Musicology) and mode (writing) (Koester 2010:68). As a result, the language contained in both Corpora is highly contextualized, reflecting Music Discourse instantiated through academic journal articles and novice essays. Such focus was selected to yield insight into this Discourse because it can reveal concentrations of patterns within it, which follow from aspects of their context (Ibid.:74).

As the purpose of this research is to identify the interactions of lexical bundles and propositional content within Music Discourse, the general design of the Corpora for this study began with selection of the written text types produced and consumed by the Music Discourse Community. In approximate order of magnitude by word count, these include: 1) reference works, textbooks and books, theses and dissertations (N.B.: These terms are used interchangeably for the master's and doctoral level between different geographic spheres of the academic world.); 2) articles, essays, lecture notes; 3) recording liner notes, program notes, music reviews and critiques, websites, blogs. Though not exhaustive, this is an extensive list, both enumerating common text types and representing a voluminous quantity of text. This overview of text types was considered in the selection of texts for a representative corpus, as 'we are only justified in claiming that a

given corpus is representative of a discourse, however we have defined it, if we have, at least in principle, access to all the texts the discourse consists of' (Teubert & Čermáková 2004:117). While it may be possible, though not practical, to include a large, representative swath of samples from all of the above text types, their sheer and continual proliferation would render such a task interminable. Moreover, not all of the above share the same context of situation. Most importantly, novices do not produce each of these types in the course of their studies. Were the entire range of types to be included in the Expert Corpus, focus on academic writing as taught to, and expected of, novices would be lost. Indeed, not all of these types are solely within the academy's domain. Furthermore, comparative balance with the Novice Corpus would be undermined both by disparity of text type and the substantial statistical manipulation required to ensure balance. Therefore, the selection of text type was narrowed for the sake of manageability and maintenance of a balanced focus within the academic register.

Given the need to maintain representativeness between the two Corpora, the choice of text type for the Expert Corpus was best determined in reference to novice production, as the scope of the latter is considerably more limited than the former. Though experts may produce any or all of the above-named text types, novices do

not. Depending on their level of study, novices may be required to write any of the above types ranging from dissertation to online texts. As the present study is focused on higher secondary and on lower-tertiary Music students, dissertations and theses were excluded. They were also disqualified for inclusion on the grounds that they represent a phase of production during which writers transition from novice to expert status, whereas the express aim of this research is a comparison between writers at two distinct stages of development. Additionally, the length of such texts and the difficulty of collecting them could well result in an imbalance of the amount of text from individual writers. Thus, the text types considered for the Novice Corpora were restricted to essays, program notes, reviews and critiques, and online content. The last three were excluded from the Novice Corpus on the grounds that online content is typically multi-modal and may exhibit a hybrid of written and spoken media. Likewise, neither program notes nor reviews are the exclusive purview of academics, and thus not pertinent to Music Discourse as defined herein; thus, they were also excluded.

Of the remaining options, essays are sufficiently generic as to be considered one of the dominant types of writing produced by novices in academic contexts. No doubt the ubiquity of this text type

is promoted by its lack of precise definition, which permits its application in various situations as an assessment instrument. Though it is necessary, and would be beneficial, to conduct detailed research into the structures and genres that define the essay within Music Discourse, such work is beyond the scope of the current study. For present purposes, a broad definition was adopted for this text type, spanning the primary socio-semiotic activities of expounding and exploring (Halliday & Matthiessen 2014:37). These two activities address the general concerns of presenting, debating and potentially even constructing knowledge that are often central to academic writing (Matthiessen 2015). Any novice text not resulting from one of these three activities was excluded in order to closely align the purposes of the texts in the Expert and Novice Corpora.

Having chosen a text type for the Novice Corpus, a suitably similar text type was then required for the Expert Corpus to maintain representativeness and uniformity of purpose. Though experts may seldom write texts titled or described as essays, other forms of writing commonly produced by them still exhibit features in common with essays. Such writings originate within academic contexts, and are therefore bound by discipline or subject; they also adhere to the primary activities of expounding and exploring. While experts may

write essays, the obligation to publish largely results in the production of articles and books, which are primarily distinguished from secondary- and tertiary-level novice essays by audience, length, and complexity of argument, and expertise.

As expounding and exploring processes are also central to the expert endeavors of book and article writing, additional criteria were needed to discern which of those expert text types are most similar to novice essays. Books were excluded from consideration on the grounds that: 1) they tend to be substantially lengthier and more extensive in scope than articles; 2) they may be more categorically specialized by purpose than articles, as is the case with biographies, studies of musical genre, catalogues and reference works; and 3) book production resides exclusively within the expert sphere, whereas articles may either be co-authored or authored by students, though it is uncertain how common an occurrence this is. Most importantly, research articles remain the primary means of disseminating and establishing knowledge (Hyland 2009), and are generally of a manageable size for triangulated analysis. Therefore, articles were chosen for the Expert Corpus because they are proportionately somewhat closer in length to novice essays (particularly extended essays) than books (though the disparity is still substantial). Although expert journal articles may not constitute

a single type, they are still suitable to the purpose of the present research as they are likely to exhibit certain patterns of functionality within Music Discourse by virtue of their register, sub-stratum within that register, and general similarities, including overall purpose and typical text length. Before proceeding with the collection of journal articles, ethics permission was obtained from the University of Nottingham Ningbo to download articles from the university's library database subscription for use in a closed corpus. The Expert Corpus built from these collected articles is held privately offline so as to avoid infringement of copyright law.

Before proceeding, discrepancies between the purpose, audience and production of expert articles and novice essays must be noted. The former are produced for dissemination to the Discourse Community for the purpose of exchanging research perspectives, while the latter are typically produced for a single reader (an expert instructor), ostensibly as a demonstration of information acquired and synthesized as knowledge, or as an exercise in discourse acquisition, which in contrast to expert writings are on occasion written hastily by a negligible number of procrastinators. These reasons notwithstanding, the ubiquity of this text type and its global dissemination, along with the range of topics addressed in such articles, meet Biber's definition of representativeness: 'the extent to

which a sample includes the full range of variability in a population'
(Biber 1993:243). While it is conceivable that articles may not
possess the full range of formulaicity inherent to Music Discourse,
the sheer volume of this text type offers greater diversity of
authorship and more capacious scope of topic than any other text
type within the Discourse, and is thus highly likely to be
representative of both lexical bundles and their ambient content
within Music Discourse. Of course, it would have been possible to
combine text types in the Corpus, such as articles and books, but
this would have posed additional challenges to balance and could
easily have overwhelmed the timeframe for this research. As to the
remaining discrepancies of purpose and audience, little can be done
to mitigate these difficulties as novices generally do not address the
expert Community in writing, nor do they often write for the same
purpose as their expert mentors. On this point, then, similitude was
accepted over exactitude, with the caveat that comparative claims
must be appropriately hedged, yet also with the recognition that
lexical bundles will be present in both Corpora by virtue of their
nature as general features of language.

3.2.2 Sampling Frame and Balance

In addition to representativeness, practicality must also determine
the size of corpora (Reppen 2010:32). Conveniently, the sample

frame for specialized corpora, such as the two in this study, can be comparatively small as the high degree of focus on a single discourse will still exhibit high-frequency lexical bundles typical of the context under consideration (Nelson 2010), provided the corpus is representative. Kennedy (1998:43) notes that specialized corpora commonly range from 100,000 to two million words, while Koester (2010:19) offers examples less than half that size. Both of these agree with L. Flowerdew (2004:67), who states: 'there is general agreement that small corpora contain up to 250,000 words'. Because the present research considers lexical bundles with their ambient content, a larger sample frame of one million words was chosen for the Expert Corpus to insure an ample range of content.

Peer-reviewed journal articles were chosen as the expert sampling unit. To insure that journals chosen were representative of the Discourse Community and its scholarly interests, articles were drawn exclusively from those ranked highly by SCImago. The SCImago Journal Rank (SJR) bibliometrics were employed because they account both for the number of citations received by a given journal and the prestige these statistics bestow on that journal; it is a measure of recognition, on the part of a given discourse community, of the value of a given journal to the community (as detailed below). Because this is a synchronic study attempting to

capture Music Discourse at this moment in time, articles were selected solely from a five-year period (2014-2018) of journal publication. Focus on the most recent expert production affords a view of the expert Discourse as it recently has been made available to novices. Because Music Discourse has yet to be investigated, a broad and balanced view of Musicology was sought through the inclusion of ethnographic, historical and theoretical studies. Conversely, journals tailored to a particular type, period, or style of music were excluded (e.g., *Journal of Seventeenth-Century Music*), as were themed issues, on the grounds that any explicit statement of specialization on the part of a journal or journal volume might compromise the generalizability of the data collected.

Generalizability was further addressed by the selection of texts from both American and British journals. This decision reflects an effort to represent Music Discourse in the broadest sense by covering the two major varieties of English found in the most prestigious literature, all of which is published either in the United Kingdom or United States. Initially, additional varieties of English were considered. To that end, journals from Australia, Canada, Ireland, New Zealand and South Africa were all considered. However, relative scarcity, cultural or subject specificity, and lower bibliometric rankings all promoted a focus on journals from the UK and the U.S. over those from other

countries. Though the exclusion of journals published in all but two countries may seem to impose limits on representativeness, such proliferation still identifies what is produced, consumed and most valued by the Discourse Community, which ultimately addresses the need for representativeness.

In contrast to specialized corpora, learner corpora, such as the Novice Corpus herein, typically have smaller sample frames, sometimes containing fewer than 100,000 words (Díez-Bedmar & Casas Pedrosa 2011, as cited in Callies 2015). Callies (2015:42) notes that the size of the corpus depends on the research, as small corpora 'are particularly useful for investigations of high-frequency phenomena at all linguistic levels'. This is fortuitous, given the frequent difficulties that may attend collection of learner data. Indeed, as will be detailed later, such collection proved challenging in this study. For the Novice Corpus, the sample frame was partially determined by the use of the Corpus for comparative purposes in a triangulated methodology and the availability of data. Unfortunately, opportunistic collection proved unavoidable in the present study for five reasons: 1) the number of Chinese students studying music in English-language schools appears to be small in comparison to other disciplines; 2) contact with students proved difficult to establish; 3) students sometimes agreed to participate yet failed to follow

through by submitting ethics forms, essays, or both; 4) some institutions staked proprietary claims on student writings; and 5) undergraduate music courses at some universities appear to be moving away from the essay as a standard measure of learner writing ability. Ultimately, learner data was collected, out of necessity, from multiple sources and participants at varying stages of secondary and tertiary study. Their texts represent the essay text type in a broad sense. As these difficulties began to multiply, I decided first to collect as much data from each participant as possible and then to select essays from each novice that would maintain the sampling frame and balance of the Corpus. Processing of the collected texts is detailed after the following the discussion of balance.

Generically, balance is understood as equal amounts of text from different kinds of sources (Hunston 2002). Since position within a single text has been shown to influence the significance of an item (Sinclair 2005), only complete journal articles were sampled for this study so as to avoid unbalancing the results with lexical bundles that might be more frequent in a given location within a text. Since it is unknown whether lexical bundles distribute evenly across Musicology articles, the use of complete texts was thus precautionary. Sinclair further notes that texts of nearly equal

length should be chosen so as to prevent specific samples from exerting undue influence over the corpus (ibid.). However, the constraints imposed by choice of text type rendered impossible a uniform distribution of tokens among the texts in the Corpus on account of variable word counts. Sampling journals of nearly equal word limits would have been preferable, but such an ideal was unattainable given the short list of publications from which to choose. The selection of these articles is discussed in Section 3.3.1 below.

Table 3.2.1 lists the word counts for all texts in the Expert Corpus by journal. As shown, the difference in total word count between the largest (220,105) and smallest (154,539) totals per journal is 65,566 words; however, most of the total word counts per journal fall close to the median of all eight journals: 184,877. This indicates that balance has been largely maintained, at least by journal.

Table 3.2.2 lists the word counts for all essays in the Novice Corpus. The secondary texts only account for approximately a quarter of the total Novice Corpus. Such a discrepancy could cause an imbalance in the results. To address this possibility, only lexical bundles found in both levels of text were retained for analysis. Furthermore, claims based on these results are appropriately hedged.

Table 3.2.1: Journal Article Word Counts

Journal	A	B	C	D	E	F	G	H
2018	10,843	8,376	8,226	7,274	9,805	15,765	7,303	9,359
2018	8,225	10,440	9,449	9,072	10,873	2,923	8,352	13,316
2018	8,490	9,151	12,845	9,084	14,009	11,325	8,334	2,922
2018	10,189	9,582	14,580	7,784	4,276	11,325	9,922	12,626
2017	9,828	9,907	10,062	9,443	11,407	7,396	7,684	3,711
2017	10,745	9,779	10,440	8,958	13,493	7,761	4,910	9,808
2017	6,660	9,021	10,763	8,794	13,954	7,316	13,152	,9776
2017	10,988	13,304	8,790	7,603	9,389	7,491	10,485	21,149
2016	10,080	11,299	8,592	10,140	19,387	10,847	4,109	3,690
2016	11,255	10,915	5,608	7,457	11,149	8,640	5,601	10,796
2016	8,842	6,204	6,671	4,397	9,158	2,740	8,566	11,921
2016	5,712	10,738	7,177	10,900	7,319	6,246	5,870	6,064
2015	5,229	7,896	7,855	9,084	13,064	12,268	8,994	7,886
2015	7,267	5,609	12,289	9,478	8,789	11,555	6,945	7,581
2015	10,631	11,555	11,273	8,177	12,692	15,941	7,757	12,335
2015	8,173	10,315	10,262	8,870	8,907	5,420	9,198	7,721
2014	10,383	10,755	10,550	8,432	7,857	8,316	9,535	7,135
2014	10,990	7,848	9,259	8,860	14,422	10,007	5,743	10,388
2014	8,892	11,984	7,061	9,888	8,699	7,912	6,805	7,214
2014	9,882	7,598	10,221	9,033	11,456	12,654	5,274	12,482
Totals	183,304	192,276	191,973	172,728	220,105	183,848	154,539	180,240

3.2.3 Frequency and Dispersion

The use of frequency to identify formulaic items stems from Firth's (1957, cited in Sinclair et al. 2004a; Wood 2015) work in the 1950s and is often used to determine which linguistic items should be

Table 3.2.2: Novice Essay Word Counts

Secondary A	1,171	Tertiary A	1,755
Secondary B	1,100	Tertiary B	1,585
Secondary C	707	Tertiary C	1,939
Secondary D	1,190	Tertiary D	2,339
Secondary E	870	Tertiary E	1,935
Secondary F	1,001	Tertiary F	2,238
Secondary G	772	Tertiary G	2,335
Secondary H	831	Tertiary H	2,350
		Tertiary I	2,290
		Tertiary J	2,040
TOTALS	7,642		20,806
GRAND TOTAL	28,448		

incorporated into teaching materials (Frankenberg-Garcia 2016). Such items may range from individual words to strings of words. A variety of extraction measures for such items may be combined for greater accuracy of identification; however, not all possible measures should be simultaneously employed, as some will prove unsuited to the Corpus under investigation, while others will be redundant (Antoch et al. 2013). For instance, measures relevant to large general corpora are not always well suited to small specialized ones, though smaller corpora may reliably employ frequency to measure significance (Evison 2010). Given the relatively small sizes of the Corpora herein, simple frequency was employed as a standard measure to identify lexical bundles within Music Discourse. This stance was adopted on the grounds that the present research

provides the first overview of lexical bundles in this Discourse and thus generality and comprehensiveness are most desirable. Indeed, the use of simple frequency as a reliable method of extracting bundles is well-suited to this project precisely because it includes 3-word bundles. Given the objectives of this research, neither the expert nor Novice Corpus was tagged, as such an operation would have been superfluous to the goal of mining for bundles by frequency.

When using frequency to mine for formulaic items, a cut-off must be chosen for the number of occurrences in the corpus and for dispersion, which is the number of texts in which a bundle appears across the corpus (Biber et al. 1999). Together, these cut-offs generate a list of Types, each of which consists of numerous Tokens, the total number of instances of each bundle in the corpus, which are presented in concordance lines. These cut-offs are set arbitrarily for the purpose of delimiting the sheer volume of data to be analyzed and are adjusted according to a study's proposed scope (Wray & Perkins 2000). Still, they must remain sufficiently generous so as to offer a broad view. Frequency cut-offs are given as the number of occurrences per million, typically 20-25 times per million, while those for dispersion are given as a minimum number of texts in which a formulaic sequence must occur to be counted, often three

to five texts (Hunston 2002; McEnery 2004; Sinclair 2005; McEnery et al. 2006; McEnery & Hardie 2012; Crawford & Csomay 2016). Furthermore, frequencies normalized to a common base (McEnery et al. 2006) can be applied to large written corpora, often 20-40 instances per million words, when making comparisons to smaller corpora, while raw cut-offs may be used for smaller corpora (McEnery 2004; Evison 2010; McEnery & Hardie 2012). The normalized frequency chosen may also be calculated to equally represent each Corpus under consideration. For instance, *The Longman Grammar* employed a low frequency cut-off of ten words per million for 4-word sequences (Biber et al. 1999). By contrast, Chen and Baker (2010) chose a minimum frequency of 25 words per million, yet followed Biber et al. (1999) in setting a dispersion cut-off between three and five texts for 4-word sequences. They took a length of four words as the typical unit of research because they are: A) common in writing; B) of a manageable length; and C) often contain other, smaller bundles (ibid. 2010:32). Similarly, other studies have focused on 4-word bundles (Wray & Perkins 2000; Chen & Baker 2016).

Though many studies have focused on 4-word bundles, often to manage the size of the data, 3-word bundles are so plentiful as to indicate that they are yet more foundational to a given discourse,

particularly disciplinary-bound bundles, as also proposed in the *Longman Grammar* (1999). Biber et al. (ibid.) observed an inverse proportion between increasing numbers of words and decreasing orders of magnitude of lexical bundles; thus, 3-word bundles are more common in academic prose than 4-word bundles by at least a factor of 10. Also, research indicating that lexical bundles are specific to a given academic register (Wray & Perkins 2000) invites speculation regarding the frequency, function and content of 3-word bundles. Moreover, the inclusion of disciplinary terms in 3-word bundles and the interaction of such bundles with ambient content is of particular interest, since an understanding of these intersections could potentially demonstrate hitherto unknown linguistic patterns that are specific to Music Discourse. Although frequency may indicate the importance of these bundles within the Discourse, it does not explain what these bundles demonstrate about the Discourse. Thus, it seems reasonable to investigate whether they are more semantically tethered to a given academic discipline than longer bundles. These questions were addressed during the research by performing a systemic functional analysis of the interaction between bundles and ambient content. This stage of the analysis is detailed in Section 3.5.3 below.

Following the convention named above, a frequency cut-off of 20 occurrences per million words was initially employed. Unlike the above studies, however, the present research had to depart from this value to meet the objective of offering the broadest possible view of Music Discourse while still retaining a manageable number of returns for subsequent functional analysis. Thus, the frequency cut-off was doubled to 40 instances across the Corpus. This was necessary given that 3-word bundles are more common by a factor of 10 than 4-word ones (Biber et al. 1999). Similarly, the dispersion cut-off for the Expert Corpus was set at 20, which is more than six and a half times greater than a typical minimum dispersion cut-off of three when mining for 4-word bundles (Hunston 2002; McEnery 2004). This dispersion value meant that a given lexical bundle had to occur across 12.5% of the texts in the Corpus before it qualified for analysis. This number ensured that even were a disciplinary-specific bundle to occur frequently in a single journal, let alone one text, it would be excluded from the data. Similarly, such a high cut-off made it less likely that bundles more commonly used in a particular sub-discipline, such as Music Theory, would be included in the data. These cut-offs resulted in 556 lexical bundle Types of 3-, 4- and 5-word length, totaling 44,700 Tokens.

Considering the limited size of the Novice Corpus, raw cut-offs for frequency and dispersion were a necessity. This is an established practice within Corpus Linguistics, which permits small corpora frequency cut-offs as low as two occurrences (Chen & Baker 2010; Chen & Baker 2016; De Cock 1998). The size of the present Novice Corpus noticeably limited the number of cut-offs possible. Likewise, a low dispersion cut-off was necessitated by the small number of texts in the Corpus. However, the fact that multiple writers contributed two articles each suggested that this cut-off be set at a minimum of three texts to prevent undue influence by a single author. Having thus determined an ideal minimum, the Novice Corpus was tested with various combinations of a 3-text dispersion cut-off and every frequency cut-off from two occurrences to ten. This process of elimination revealed that a frequency cut-off of four, combined with a dispersion cut-off of three, yielded the largest number of bundles with the least amount of overlap among Types for a total of 172 Types of 1,220 Tokens. This expedient was commended by first having calculated all overlapping bundles in the Expert Corpus. This process eliminated the problems of manually calculating and removing all bundles rendered redundant through overlap, and of omitting those several low-frequency bundles that were the product of a specific essay question. The latter proved problematic because the four secondary writers each submitted two

essays, and each of those were answering a specific essay question. This resulted in some bundles appearing that were obviously connected to the essay question. Having thus mined the Corpus, these 172 Types were compared with the secondary and tertiary student subcorpora to determine if any should be excluded on the basis that they were responses to a single assignment from the secondary texts.

3.3 Data Collection

3.3.1 Expert Journal Articles

Considering the aforementioned criteria for representativeness and sampling frame and the wish to include major varieties of English, a total of 160 articles from eight journals across five years were compiled for the Expert Corpus. These were drawn equally from four American and four British journals, with 20 articles selected from each. The number of journals sampled resulted from the intersection of a sampling frame target of one million words with the following three criteria: 1) Bibliometrics Ranking, 2) General Focus, 3) Distribution by Variety of English.

Since it ranks journals according to their prestige within the academic community, both by number of citations and countries where the citations are being made, SCImago was chosen as the

bibliometrics criterion for the Expert Corpus (SCImago, n.d.). Taking both of these factors into account helped ensure that the journals chosen are not only prestigious globally but also within their own countries. This is of further importance for L2 novices and their instructors as the perceived prestige of English-speaking countries among Chinese students can contribute to the overall impression that journals produced and cited in those countries are most valuable.

SCImago (SJR) lists journals according to frequency of citation, and is thus an indicator of the relative value afforded a journal by the academic community. Given their prestige, these journals are those most likely to be generally consumed by the relevant expert community with the highest frequency. Initially, the first 50 music journals in were considered as candidates for inclusion in the Expert Corpus.

However, SCImago could not serve as the sole criterion for journal selection, as the majority of the 50 most highly ranked journals focus on narrow fields within music-related studies. Such taut focus would have defeated any attempt to obtain a broad view of Music Discourse. Thus, the second criterion of generalizability was met by categorizing the focus of each journal in the SJR list, as determined

both by its SJR description (see Appendix 1) and its statement of purpose taken from each journal's website (see Appendix 2 for the complete list of Journal Statements of Purpose), and then eliminating those of narrow scope. This resulted, for instance, in the highest ranked journal, IEEE Signal Processing Magazine, being eliminated on the basis that it is focused on technological issues. Similarly, all Music journals dedicated to acoustics, aesthetics, audio and technology, education, library science, mathematics, perception and psychology, or therapy were excluded. Also rejected were all journals dedicated to a particular geography or culture, (e.g., *Journal of African Cultural Studies*), a given style (e.g., *Popular Music and Society*), an era or time frame (e.g., *Twentieth-Century Music*).

This process of elimination rendered three types of journals dedicated to Musicology, Music Theory and Ethnomusicology, as self-defined on their respective websites (see Appendix 3). In the broadest sense, Theory and Ethnomusicology fall under the umbrella of Musicology, which developed from a nineteenth-century interest in the description and cataloguing of musical works, and later expanded to a twentieth-century interest in music as socially-situated phenomenon; both conceptions retaining a focus on music as object (Hitchcock & Deaville 2013). Though they appear to

represent somewhat narrower focuses within Musicology, which would seem to commend their exclusion from the Expert Corpus, the analytical approaches of Music Theory and Ethnomusicology may be applied to music of any style, era and culture, as well as any physical phenomena related to music as performance, artifact or practice. Therefore, they share a breadth of scope with Musicology and the general purpose of investigating music-as-phenomenon, again as evidenced by their online descriptions (Appendix 3). As an example, the website for Music Theory Spectrum states that it is:

> A leading journal in the field and an official publication of the Society for Music Theory, Music Theory Spectrum features articles on a wide range of topics in music theory and analysis, including aesthetics, critical theory and hermeneutics, history of theory, post-tonal theory, linear analysis, rhythm, music cognition, and the analysis of popular musics. The journal welcomes interdisciplinary articles revealing intersections with topics in other fields such as ethnomusicology, mathematics, musicology, philosophy, psychology, and performance. (Society for Music Theory, 2018. Home page.)

Of particular note here is the reference to scholarship that intersects with other fields. Lastly, it should be noted that exclusion of Music Theory and Ethnomusicology journals would have unduly constricted the range of journals included in the Corpus, rendering it difficult to make claims about Music Discourse generally, as opposed to a potentially narrow definition of Musicology specifically. This is a danger given that the label *Musicology* is applied to music scholarship in general outside of the United States, whereas American music scholars distinguish it from Music Theory (Duckles et al. 2001). This fact was taken into account when choosing U.S. publications and addressed by selecting Musicology journals whose self-description invited submissions from Ethnomusicologists and Theorists.

Though a larger number of journals was initially sought in order to reach a one-million-word target, by drawing upon the widest possible range of journals among those most highly ranked, the nexus of the above criteria winnowed the number of journals to nine from among the top 50 SCImago rankings. Of those nine, only three were British, and thus consideration was extended to the top 60 SCImago rankings to gather four British and four American journals of sufficiently general focus to make an even total of eight. The decision to include equal numbers of articles from UK and U.S.

publishers was made in an attempt to avoid privileging one variety of English over another. Though country of publication does not necessarily correlate with nationality of authorship, the publication criteria favored by these various publishers may nevertheless reflect some of the differences among these varieties of English, as evidenced by their editing protocols. Moreover, journals with national associations (e.g., *Journal of the Royal Musical Association*) might enjoy a higher percentage of domestic contributors as such writers may be members who read and privilege their association's journal over others. Lacking membership statistics, these criteria were applied out of an abundance of caution for the maintenance of representativeness and balance. As previously mentioned, it was not possible to sample journals from English-speaking countries other than England and the United States, as none had a sufficiently high SJR ranking.

Below, then, are the eight journals ultimately selected for the Expert Corpus, listed in Table 3.3.1 by country of publication, sub-discipline, and SJR number and title. In Table 3.3.2, the same journals are listed by their SCImago ranking. Interestingly, none of them are drawn from the top 20 rankings, as all of those journals were focused on specializations, primarily music education and technology, which may have ramifications for future study.

Table 3.3.1: Journals by Country and sub-Discipline

Country	Musicology		Ethno-musicology	Theory
UK	45 Journal of the Royal Musical Association	48 Journal of Musicological Research	36 Ethno-musicology Forum	56 Music Analysis
U.S.	47 Musical Quarterly	35 Journal of Musicology	29 Ethno-musicology	28 Music Theory Spectrum

Table 3.3.2: SCImago Journal Rankings

SJR	Journal	Country
28	Music Theory Spectrum	U.S.
29	Ethnomusicology	U.S.
35	Journal of Musicology	U.S.
36	Ethnomusicology Forum	UK
45	Journal of the Royal Musical Association	UK
47	Musical Quarterly	U.S.
48	Journal of Musicological Research	UK
56	Music Analysis	UK

From each journal, four articles were taken for each year from 2014-2018 for a total of 20 articles from each journal. Only original articles were chosen. As previously mentioned, themed issues were avoided so as not to imbalance the Corpus. Thus, the first issue in 2018 of Music Analysis was rejected as it was titled 'Special Issue on

Film Music' (Wiley Online Library: Music Analysis [WWW Document], n.d.). Aside from this particular consideration, however, articles were simply taken in consecutive numerical order as they appeared in each journal (i.e., the first four articles), starting from the first issue of each year that was not themed. No attempt was made to impose further selection criteria as this could potentially have introduced bias in the choice of articles. This decision partially accounts for the fact that articles of varying lengths were selected. While it may have been possible to choose articles non-consecutively from some journals to achieve a greater balance in the number of Tokens for that journal, not all of these journals contained a sufficient number of articles per year to render this a consistently reliable method of selection.

As a result of the above selection criteria, the number of words per article in the Expert Corpus ranged from 2,740 to 21,149, with a mean of 9,292, median of 9,118, and mode of 10,440. Three methods of addressing such a wide range were considered: 1) excise the longest and shortest articles from each journal, 2) divide the Expert Corpus into subcorpora based on stated purpose of publication for comparison against the whole Corpus, or 3) accept and acknowledge potential consequences of this disparity of text length with the justification that the large number of articles and

resulting amount of text significantly mitigates the influence of any one article; thus, viable claims about the use of lexical bundles across the Discourse can still be produced, particularly if findings from the same are consistent with earlier research. The first solution had to be discarded because some journals contained two or more articles of insufficient or excess length per year, so that excising them would have excessively complicated any attempt to achieve balance by discarding multiple articles per year. Subsequently, the second and third solutions were combined to produce hedged claims based on a comparative view of data results from the Corpus and subcorpora, organized by journal and area (i.e., Ethnomusicology, Musicology, Theory), to highlight any discrepancies in frequency between items in the entire Corpus or its parts.

Articles collected for the Expert Corpus totaled 1,486,653 words. As this exceeded the original plan to collect one million words by nearly 50%, an experimental comparison was made of the number of returns from this Corpus and a version of this Corpus reduced to some 1,100,000 words. This reduced form was generated by excluding all articles from 2014. Upon mining both forms of the Corpus, with frequencies and dispersion proportionally adjusted for the differences in size, it was discovered that the smaller version contained only some ten lexical bundle Types fewer than the larger

version yet some 10,000 more Tokens. This small difference in the number of bundles, combined with the substantial increase in number of Tokens, prompted the use of the larger, original version of the Corpus, to obtain a somewhat more comprehensive view of lexical bundles in the Discourse while simultaneously reducing the burden of manual analysis.

3.3.2 Novice Essays

Data collection for the Novice Corpus proved more challenging than anticipated, as it was hindered by a number of factors. Potential difficulties were identified at the outset of this project and contingencies plotted to avoid disruptions of, and circumvent obstacles to, data collection. Despite these preparations, multiple set-backs impeded progress, ultimately resulting in a smaller number of texts and overall word count than initially targeted. Nonetheless, the data obtained proved useable for reasons discussed below.

A refinement process similar to that delineated for the Expert Corpus was applied to the selection of texts for, and construction of, the Novice Corpus. As stated above, such specialized corpora are generally small. This is often by necessity as the availability of data for such corpora may be constrained, as was the case here. Thus,

an initial target of 200,000 words was set for the Novice Corpus. Given the specialization of both Expert and Novice Corpora, this number would have been more than sufficient to exhibit high-frequency formulae in the texts, while simultaneously limiting the resulting data to a manageable amount. Gathering said amount of novice text proved untenable, however.

Because articles do not constitute a significant portion of novice output, a different type of text that reflects the requirements of novice writing was collected for this Corpus. In order to study novice writings from Chinese music students, my original proposal was to collect samples of International Baccalaureate (IB) Diploma Programme Extended Essays from IB Centres throughout China. The stated aim of the IB Diploma Programme is to prepare students for disciplinary research and writing at the tertiary level; the Extended Essay, which is required in the final year of IB study, is intended to assist students with this preparation (International Baccalaureate 2018b). IB Extended Essays are produced by novices in their final year of secondary studies, just as they prepare to enter university. The advantages of collecting Extended Essays included: uniformity of essay size (4,000-word limit) and general prompt (requires analytical description of music); a shared purpose of preparing students for university study (ibid.); the substantial number of IB

Centres in China with music students (60); the added opportunity of studying writings from novices for whom English is an L2. This last point would have added value to the present project by directly connecting it to the areas of EAP and SLA research, thus broadening its scope. As China already hosted more than 60 IB Centres at the time of collection that offered music as a subject, this plan potentially could have produced a large data sample consisting of uniform text type and length produced by Chinese novices.

Nonetheless, this proposal proved impossible. Though an initial inquiry with the IB Regional Headquarters in Singapore yielded a positive response to the request to collect marked IB Extended Essays, the same inquiry to IB International Headquarters in Cardiff, England, was answered with a proprietary claim for all marked Extended Essays, thus blocking such data collection. Rather than attempt to collect Extended Essays prior to their marking, which would have presented logistical and analytical difficulties, and in anticipation of such a response from the IB Programme, five alternate collection plans had been devised: 1) collect marked IB essays that had been class assignments, rather than Extended Essays, and were therefore not proprietary; 2) collect essays from students at international schools in China or from domestic music schools, such as Tianjin Juilliard or the Shanghai Conservatory; 3)

collect essays from International Baccalaureate Centres abroad should collection in China prove problematic; 4) collect texts from Chinese music students at foreign high schools or universities; 5) reorient this research to a diachronic study of expert writings. In reality, data collection required the aggregation of aspects of the above plans, particularly given the reluctance of novices to share work they considered of inferior quality.

Regarding the first plan, inquiries were sent via e-mail to all IB Diploma Programme Centres in China that offer music. Of these, a single school in Beijing offered to contribute eight complete texts, though the response only came eight months after the initial request. These essays were collected after obtaining ethics consent. They represent the work of four novice writers, all secondary students, writing on two topics assigned as classwork. After preparation for inclusion in the Corpus, these eight texts totaled 7,746 words. Given the obstacles encountered with the IB Programme and Centres, it was decided to forego any attempt at the third plan, which was also IB-oriented.

Next, the second plan was considered, since the possibility had then arisen of collecting class or application essays from the first cohort of graduate students to matriculate at The Juilliard School's new

campus in Tianjin, P.R.C. The immediate problem with this plan was the level of the students. Collecting essays from master's students significantly surpassed the original proposal to study writings from higher secondary-level and lower-tertiary level students. As with the proposed IB data collection, however, this plan enjoyed the merits of uniformity of text length and purpose, though text type may have been less predictable given the absence of a prompt. Unfortunately, this opportunity was ultimately voided by an administrative decision to delay the first cohort's matriculation by an entire academic year. At nearly the same time, one of the assistant dean's at Tianjin Juilliard, who had previously taught at the Manhattan School of Music, offered a large collection of writings from former music students at the latter school. Unfortunately, upon requesting ethics consent, this offer was rescinded.

Subsequently for the second plan, novice writings were sought from the Pre-College Division of the Tianjin Juilliard School. At the time, this seemed fertile ground for collection as I was in contract negotiations with the school and was assured that data collection would be a simple matter, though it was unsure how much text might be available or how many secondary students would be enrolled during the first year of Juilliard's Pre-College Program. After being hired, I contacted my colleague at Juilliard New York, who is

Director of their ESL program, to request data collection from Chinese students currently studying there in order to enlarge the Novice Corpus. I was again assured that collecting from the New York campus would be a simple matter, as similar requests were routinely submitted to the school, reportedly with positive results. Nevertheless, after several months of further inquiry, I was informed that Juilliard has a long-standing policy prohibiting data collection. At this juncture, it was increasingly apparent that additional data collection would prove highly problematic. These unsuccessful attempts, combined with the time constraints of the research, finally indicated that some form of convenience sampling was unavoidable.

Not wishing to abandon a focus on Chinese novice writers, I turned to the fourth option of collecting essays from students studying abroad. This took the form of contacting Chinese undergraduates studying music in English-as-Medium-of-Instruction (EMI) universities and requesting marked examples of their written essays. This was less systematic than the approach taken with the IB Centres, as it involved contacting colleagues to ask if any of their former students might be willing to offer writing samples. In this manner, texts were ultimately collected from eleven additional students studying in Australia, England, and the United States. A single text was collected from each of two of the students, two texts

each from two of the students, and a trove of six texts from a single student. Regrettably, such a drastic imbalance in the number of texts from individual writers made it necessary to eliminate the two largest and two smallest texts from the set of six contributed by a single student, thus leaving two texts. Altogether then, the texts from these five students were added to those of the IB students for a total Corpus of 28,448 words.

The entire process of collecting this novice data spanned a period of nearly two years and proceeded sporadically. Under normal circumstances, it might have been possible to collect additional texts by requesting that participants in this research invite their classmates to submit essays. In fact, I made such requests to all of the tertiary students who contributed texts, in some instances asking twice at different times. Because most of this data collection occurred over the same period as the COVID-19 pandemic, however, these student participants were all studying online and thus had no personal contact with classmates. Even the few of them who knew there were other Chinese students in their online courses were understandably reluctant to approach strangers via e-mail or social media simply to pass along a research request.

As a consequence of this situation and these obstacles, I finally closed the Novice Corpus to further additions, thus limiting it to the near-30,000-word count to maintain the balance achieved by excluding certain collected texts. Still, this limited size is defensible for the following reasons. Firstly, specialized corpora often consist of a limited number of words as they target specific information within a limited scope. Secondly, claims based on findings from this Corpus are sufficiently hedged. Thirdly, data specific to essay prompts are expurgated to render remaining returns generalizable. Fourthly, these remaining returns exhibit a high degree of correspondence with those from the Expert Corpus, thus further supporting the generalizability of the data. Fifthly, generalizable bundles exist in the Novice Corpus in nearly the proportion identified by the Longman Grammar, thus further supporting the generalizability of the data. Lastly, limitations of this research are acknowledged alongside recommendations for further research.

Though considerably smaller than the initial target, the number of words in the Novice Corpus still constitutes a balanced data set from which observations and cautious extrapolations may be drawn. The fact that students studying in various countries are included is a welcome addition, given the division of the Expert Corpus between American and British forms of English. Though there is no standard

minimum limit for the number of Tokens in a specialized learner corpus, there is general agreement that such corpora can nonetheless prove valuable, even with numbers ranging from 10,000 to several tens of thousands of words, provided they can still address the research questions at hand (Biber 1990; Bowker & Pearson 2002; Carter & McCarthy 1995; Dutra, Orfanó & Almeida 2019; Gesuato 2011; Vyatkina, Hirschmann & Golcher 2015). Thus, more significant than the minimum number of Tokens is the purpose of this Novice Corpus, which here is intended for comparison with the Expert Corpus through a detailed analysis of how competently novices connect lexical bundles, both generally academic and disciplinary-specific, to propositional content. The level of detail required by such an analysis makes the number of Tokens less important than the shared purpose of the texts. The greater concern here is the fact that the eight IB essays were written on only two topics. This is unlikely to negatively impact the inclusion and use of lexical bundles, given their pervasiveness; yet it did prove troublesome for setting a dispersion value for the Corpus. As with the Expert Corpus, this problem was addressed by setting a relatively high dispersion cut-off, and by dividing the Corpus into subcorpora, by secondary and tertiary levels that could be compared with each other. This follows Gilquin's (2015) admonition to consider

the level of novices' development, including their education or proficiency at the time of data collection.

As noted earlier, ethics consent was obtained from the University of Nottingham Ningbo before proceeding with the above data collection, and all work was undertaken transparently. Students were asked to read explanations, provided in Chinese and English, of the proposed research and use of their writings. Everyone who was asked kindly agreed to participate and signed consent forms stating permission for use of their writings (Appendix 1: Ethics Consent). All novice texts were collected in electronic format by e-mail or social media to facilitate processing for the Corpus, but were anonymized before entry into the Novice Corpus by the removal of name, date and institutional information, including assignment rubric. The anonymized texts were then converted to PlainText format using the software AntFileConverter (Anthony 2017). This program eliminates scores, graphs, tables, charts and diagrams, but sometimes replaces them with strings of code that had to be removed. Given the small size of this Corpus, it was decided that normalization of orthography and spelling, including the correction of spelling errors, was necessary to insure that no lexical bundles would be inadvertently excluded when processed in AntConc. Finally, both the original files and the Corpus have been stored privately

offline to ensure anonymity. Appendix 5 lists institutions and general data for contributing novices, as well as the date of collection and title for each file. No other information has been stored, save that regarding schools and countries where novices were studying (see Appendix 3).

3.4 Data Processing

3.4.1 Overview

Both Corpora were processed using Laurence Anthony's free concordancing and text analysis software, AntConc (Anthony 2018). This software enables searches for n-grams (i.e., lexical bundles) with cut-offs for frequency and range (i.e., dispersion) and a file view function that permits researchers to see the relative location of Tokens in each text within the Corpus (ibid.). The prepared files were loaded into AntConc (Anthony 2018), word lists were generated, and then Cluster/N-Grams (i.e., lexical bundles) of three, four and five words were searched using a total frequency cut-off of 40 for the entire Corpus. Range was set for 20 texts. This relatively lower range was chosen because bundles are somewhat less frequent in the academic written register (Biber et al. 1999). Lastly, the results generated for lexical bundles, concordance lines and concordance plots were cloned and saved as both Apple Pages

documents, Plain Text files, and Numbers spreadsheets listing order, frequency and range for each lexical bundle Type and Token.

Having selected all of the articles, they were first filed as PDFs according to Sinclair's taxonomy to insure that all concerns of representativeness had been met for the Expert Corpus: mode of text (e.g., writing), type of text (e.g., journal article), domain (e.g., academia), language (English), location (e.g., originated in U.S.), and date (2005:7). Subsequently, the body of each article was copied from its PDF to a MacBook Pages file. This proved the most expedient method of isolating the text by avoiding the inclusion of time or download stamps, publishers' information, titles, authors, abstracts, notes, music examples, and bibliographies. These files were labelled generically according to SJR number, year and order of selection within that year, and word count. Hence, C. Catherine Losada's 'Complex Multiplication, Structure, and Process: Harmony and Form in Boulez's Structures II', which was the first article collected from Music Theory Spectrum (2014), was filed simply as 28-4a 10843. Each file was then converted to Plain Text format, retaining the same file name.

Regarding concordance lines, AntConc was set to retain 110 characters in both the preceding and succeeding slots for each

bundle Token so as to include a sufficient amount of ambient content for analysis. This was necessitated by the presence of lengthy nominal groups including multiple post-modifying prepositional phrases. One difficulty in retrieving this ambient content is that AntConc counts both punctuation and symbols among letters as part of these 110 characters, resulting in some Tokens that are bisected in ways that render them useless for analysis (e.g., by a comma). Furthermore, the limit of 110 meant that open and close quotation marks were sometimes beyond this boundary, thus obscuring portions of text that are actually part of lengthy quotations. These problems were addressed through manual analysis of all concordance lines.

AntConc can be set to recognize numerals, but this results in all numerals being read, including those for endnotes. As these are typically positioned next to words or punctuation without spaces after rich text files have been converted to plain text (e.g., superscript numerals are changed to normal script), setting AntConc to recognize these would negatively impact the software's recognition of words. Thus, the software was set to ignore numerals, which resulted in certain Types or Tokens being interrupted by numbers. These then had to be manually removed.

3.4.2 Overlapping Lexical Bundles

Following Chen and Baker (2016), overlapping bundles that could be aggregated into longer strings were identified through manual calculation to reduce the research burden. This was only done for bundles completely subsumed by longer ones. For example, Table 3.4.1 lists all the bundles from the Expert Corpus that contain the word 'end'. The first two bundles of three words are aggregated until

Table 3.4.1: Lexical Bundles containing 'end'

Rank	Frequency	Range	Lexical Bundle
4	439	119	the end of
8	294	107	end of the
12	259	99	the end of the
26	263	88	at the end
30	226	85	at the end of
76	122	63	at the end of the

the 4-word bundle ranking 12th. The same aggregation occurs with the fourth and fifth bundles in Table 3.4.1, which also are combined in the final 5-word bundle. In cases where only some Tokens of a lexical bundle are subsumed under a longer one, those that are not became the sole Tokens considered when retrieving ambient content.

3.5 Data Analysis

3.5.1 Overview

Analysis of the data, both lexical bundles and ambient content, from the Expert and Novice Corpora was analyzed in eight stages, some of which had not been initially planned but which were suggested by findings from other stages. These eight are: 1) word class analysis; 2) word function analysis; 3) analysis of prepositions and phrases; 4) extraction of disciplinary terms, either musical terms or related nominals; 5) analysis of global views of complete texts; 6) bundle discoursal functions; 7) mapping of ambient content; 8) comparative analysis of findings from both Corpora. Each of these is detailed below.

Word classes were analyzed for all bundle words. This involved assigning a part of speech to each word or, in some cases, to groups of words. The latter include compound and complex items, such as 'a and b' and 'as well as'. Any uncertainty regarding compounds or complexes was resolved by consulting the *Cambridge Grammar of English* (Carter & McCarthy 2006) and Halliday and Matthiessen's (2014) *Halliday's Introduction to Functional Grammar*. This labelling enabled the categorization of lexical bundles into syntagmatic Types, which represents a first step toward understanding the recurrent structures that underpin the Discourse. The resulting Types,

especially few in number for 4- and 5-word bundles, could be of use to disciplinary writing instructors who need a summary of instrumental grammatical patterns in the Discourse.

Next, word functions were analyzed to determine any systemic patterns. This analysis relied solely on Halliday and Matthiessen's (2014) explication of the nominal group and the word functions that constitute it. For instance, this stage not only identified nominals but more specifically revealed the presence of several numeratives. Following on this discovery, prepositions were analyzed to identify post-modifying phrases in nominal groups that could contain a Thing dissociated from the Head of the nominal group. This structural relationship is represented by the preposition 'of' (Ibid.:394), so bundles containing it were analyzed separately. This revealed a substantial number of extended numeratives, which were analyzed according to Halliday and Matthiessen's (Ibid.:395) categorization of them.

Then, disciplinary terms in the bundles were identified, separated from non-disciplinary ones, and organized into categories that could be mapped graphically. This was accomplished by representing their relative frequencies as differing font sizes in a map. At the same time, these facets of music were then represented as a constellation

of relationships. The resulting map displays the disciplinary concerns of the Discourse as found in the lexical bundles.

Having divided bundles into disciplinary and non-disciplinary categories, complete texts from both Corpora were selected for the creation of global views. This was accomplished by assigning two different colors to these two types of bundles (i.e., disciplinary and functional, or purely grammatical) and highlighting all instances of them in the texts. Next, structural sections of the text were also assigned colors used to highlight three label headings: Introduction, Body, Summary/Conclusion. Quotations were highlighted in grey to show parts of a text that did not originate from the author. Finally, the font color was changed to white and its size reduced to 3, which permitted even the selected journal articles to fit on a single page. The choice of white font also renders all of the text invisible except the highlighted bundles and quotations. This avoids any infringement of copyright as only a minuscule portion of each text is given. Altogether, this generated a global view of lexical bundles across a text, divided into disciplinary and functional, and their interaction within a text to reveal their relationships to one another, their locations, and relative proportions.

The final step in analyzing the bundles themselves involved the application of Biber et al.'s (2004) taxonomy of bundle discourse functions. Having determined each function, it became possible to identify referential bundles as most likely to connect to ambient nominal content. Initially, I had planned to apply this taxonomy at its most delicate level, identifying not only each major category of bundle function but also all sub-types. Once the work of categorizing began, however, it became apparent that that level of analytical delicacy would not further the purposes of the current study since the only two reasons to apply this taxonomy were to measure the similitude of bundle Types in the Discourse against Biber et al.'s findings (ibid.) and to identify referential bundles capable of connecting to ambient propositional content.

Identification of referential bundles helped somewhat reduce the number of ambient slots that needed to be read and analyzed in each concordance line, though not greatly. All Tokens of all lexical bundles were placed into the central column of a spreadsheet. The preceding and succeeding 110 characters of each concordance line were then placed in the cells adjacent to this column. Nominal groups or portions thereof were then identified from these adjacent cells, or slots, for all referential bundles and copied into the next outer adjacent cells of each row. From these, words functioning as

numeratives, classifiers, nominals, and post-modifying prepositional phrases were extracted and placed in the next outer adjacent cells. Finally, the Thing was identified in each slot and placed alone in the final outer adjacent cells. In cases where a pronoun was substituting for the Thing, the concordance line was analyzed to determine which Thing was referenced by the pronoun and this was placed in the outermost cells to avoid undue influence from pronominal substitution.

After the exhaustive spreadsheet analysis of ambient content, all Things identified in those outermost cells were arranged in alphabetical order for the purpose of creating a typology that would account for these findings. I devised this typology as a hierarchy of generalizations. Hence, 'forte' was placed in the category 'dynamics', just as 'clarinet' was placed in the category of 'instrument'. All titles of musical pieces or works were simply filed under the heading 'title'. These generalizations were then used to create a cartographic representation of the propositional categories in the Discourse. At this stage, it was sometimes necessary to determine instances of potentially polysemous items (e.g., 'note') so as to avoid categorizing them incorrectly. This necessitated careful reading of the context for each slot in order to determine the appropriate category. Lastly, findings from all of the above analytical stages of

both Corpora were compared to identify similarities and differences in lexical bundle usage and propositional content. This was accomplished simply by comparing statistics and graphics from each Corpus and noting differences in frequency, use, and location within a text.

3.5.2 Analytical Rubric for Ambient Content

The following analytical rubric was devised to maintain a narrow focus on identifying the experiential function of Head/Thing in nominal groups. It was applied during stages 4 and 7 of the analysis.

1. Lexical Bundles: Possessive Types and Tokens of bundles, such as ExCo 10 ('of the music's'), were omitted, sometimes resulting in a loss of disciplinary lexis. This was the case, for instance, with 22 Tokens of ExCo 29 ('of the song') and 19 Tokens of ExCo 32 ('of the work').

2. Ambient Content: Only directly adjacent nominal groups or portions thereof were retained for analysis in order to identify the Thing in the nominal group. Punctuation on either side of a bundle disqualified the content in that slot as it was bounded. Rankshifted relative clauses were omitted from analysis because they did not contain the Thing of the nominal group. Post-

modifiers were retained for analysis only when they were defining and had the potential to contain a dissociated Thing. Additive compound Things resulting from parataxis within the nominal group were all retained. All quotations were excluded from analysis.

The content from all of the productive slots was then separated into two groups: disciplinary specific (e.g., 'movement', 'forte', 'articulation'), and that which was not obviously disciplinary in nature, such as generic nominals (e.g., 'way', 'place'). In ambiguous instances, where a word could belong to both categories (e.g., 'time'), it was assigned to the latter category, labelled Extramusical, which required no further analysis. By contrast, all disciplinary Things were placed in a typology that organized them according to facets of music, such as rhythm, melody, harmony, dynamics, tempo, articulation, structure, and analysis, to name a few. All of these categories are presented in the next two chapters on findings.

Occasionally, consistent application of the rubric proved challenging, as was the case with an article on a bread-fruit-summoning ritual (Diettrich 2018), in which a large portion of the discussion revolves around a ritual that includes singing. Given that a substantial portion of the discussion is dedicated to a description of the ritual,

many of the Things identified in ambient slots did not meet the criteria for inclusion in the disciplinary typology, despite being remotely connected to an analysis of music in a social context.

3.5.3 Cartography of Ambient Content

Before a typology of Music Discourse could be constructed, a definition of music was needed to ground and guide it. The definition constructed for this is:

> The phenomenon of music is a socially situated activity conducted through the medium of sound that unfolds in time, the signification of which may be interpreted individually.

This definition takes into account the social context in which music arises (Nattiez 1990), the varied experiences in time of performing and listening (Kania 2017), and the idea that music lacks clearly definable signification (Gadamer 2004b, Scrutton 2016, Tuan 2009). Explicit here is the move from the universal to the individual, music's social foundation, its aural and temporal manifestations, and its potential for polysemous signification. The above allows for sonic expressions that many listeners might simply find cacophonous. A lack of euphony under a conventional guise such as melody might

146 of 460

well seem inherently nonmusical to the majority of audiences. Nevertheless, it is precisely the recognition of this aspect which permits the inclusion of all manner of musical expressions, ranging from remote world musics to post-structuralist experiments with indeterminacy. Again, the necessity of such an inclusive definition is warranted by the very breadth of the Discourse being investigated. Based on this definition, it was possible to construct a typology of general categories that accounts for all Things mined from the ambient propositional content of the lexical bundles in the Discourse, as stated above, and consequently organize them into a cartographic representation of the Discourse.

One complication to the present methodology that is unavoidable yet also difficult to solve is the overlap of bundles, which causes a very few slots to be counted and analyzed twice. For example, in LB 4 ('in which the') Token 99 elides this bundle with LB 307 ('the first of') to create: 'in which the first of'. Such overlaps were determined to be exceedingly few in number during manual analysis of the ambient slots.

4 Expert Corpus Findings

4.1 Introduction

In the discussion that follows, five views of Music Discourse are presented: 1) the structural composition of lexical bundles (hereafter LB) by word class; 2) LB word functions; 3) disciplinary content within LB; 4) discoursal functions of LB; 5) ambient content from slots surrounding LB.

This chapter contains discussion of the findings from the Expert Corpus (hereafter ExCo). Comparative analysis of these findings with those from the Novice Corpus (hereafter NoCo) is presented in the next chapter. For each Corpus, discussion begins with a brief overview of the number of LB Types, their frequency, range, and distribution by number of words. This data is then placed in the context of word counts per text (i.e., ExCo journal articles; NoCo essays). Next, syntagmatic structure and Type are analyzed by word class, followed by analysis of bundles by word function with a particular focus on nominals. Next, all prepositions contained in each Corpus are presented by frequency, after which genitive bundles (i.e., those containing 'of') are analyzed, since they form a significant portion of nominal groups within the Discourse. These LB are then divided into categories that form either numeratives and extended numeratives, or other nominals.

At this juncture, an additional research item is introduced: ExCo extended numerative bundles of 6-8 words. I added these upon discovering several examples of such LB built from overlapping shorter ones. As longer word strings, these LB form a convenient group of high-frequency bundles that could be easily memorized by novices for use in their own writing. As these only exist in the ExCo, this section does not include analysis from the NoCo. Following this, all bundles containing disciplinary terms are categorized according to their semantic content.

I then return to analysis of 3-, 4-, and 5-word LB, viewing complete texts (i.e., journal articles and essays) at the global level to highlight where disciplinary and non-disciplinary bundles appear by section within each text. Analysis then proceeds to the discoursal stratum, employing Biber et al.'s (2004) taxonomy to categorize bundle functionality within the larger discourse. Finally, all ambient nominal content of bundles (both preceding and succeeding slots) is surveyed, categorized, and mapped; high-frequency categories are also mapped by content.

4.2 LB Type, Frequency, Range

Having eliminated subsumed and overlapping bundles, the total number of remaining LB is 536. Table 4.2.1 divides these 536 Types by the number of words per bundle, from 3- to 5-word bundles.

These LB are shown as percentages of the total in Chart 4.2.1 for the 536 LB that remain after eliminating subsumed and overlapping bundles. The resulting proportion of 90:9:1 for 3-, 4-, and 5-word LB approximates the standard proportion of an order of magnitude found in *The Longman Grammar* (1999). (N.B.: The proportion herein would more closely approximate 100:10:1 had subsumed bundles not been eliminated.)

Table 4.2.1: LB Types by words per bundle

	3-LB	4-LB	5-LB
Number of Types	482	49	5

Given that the methodology of this study is built on frequency, a cursory overview of the data can be gained by viewing LB that comprise the most frequent 10% of all Types. Among these, only a few contain words readily identifiable as disciplinary-specific terms. One group of terms, largely numbers, seem to belong to the

Chart 4.2.1: LB by words per bundle
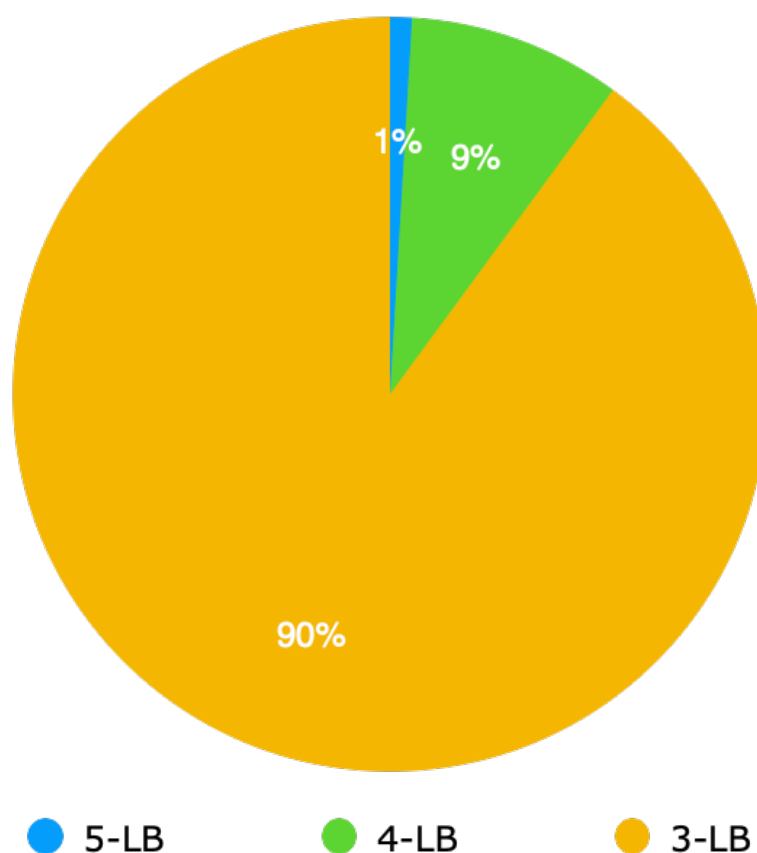
Discourse as part of the discussion of scores. Finally, there are bundles that contain no discernible disciplinary terms. While these bundles are analyzed in detail later, their proportions are represented below in Chart 4.2.2.

Among the most frequent 10% of all LB, those that definitively include disciplinary terms are shown in Table 4.2.2 below. Six of the seven bundles refer to music generally, or to a piece of music, or a portion thereof. The seventh may not refer to something musical yet

Chart 4.2.2: Terms among top 10% of LB
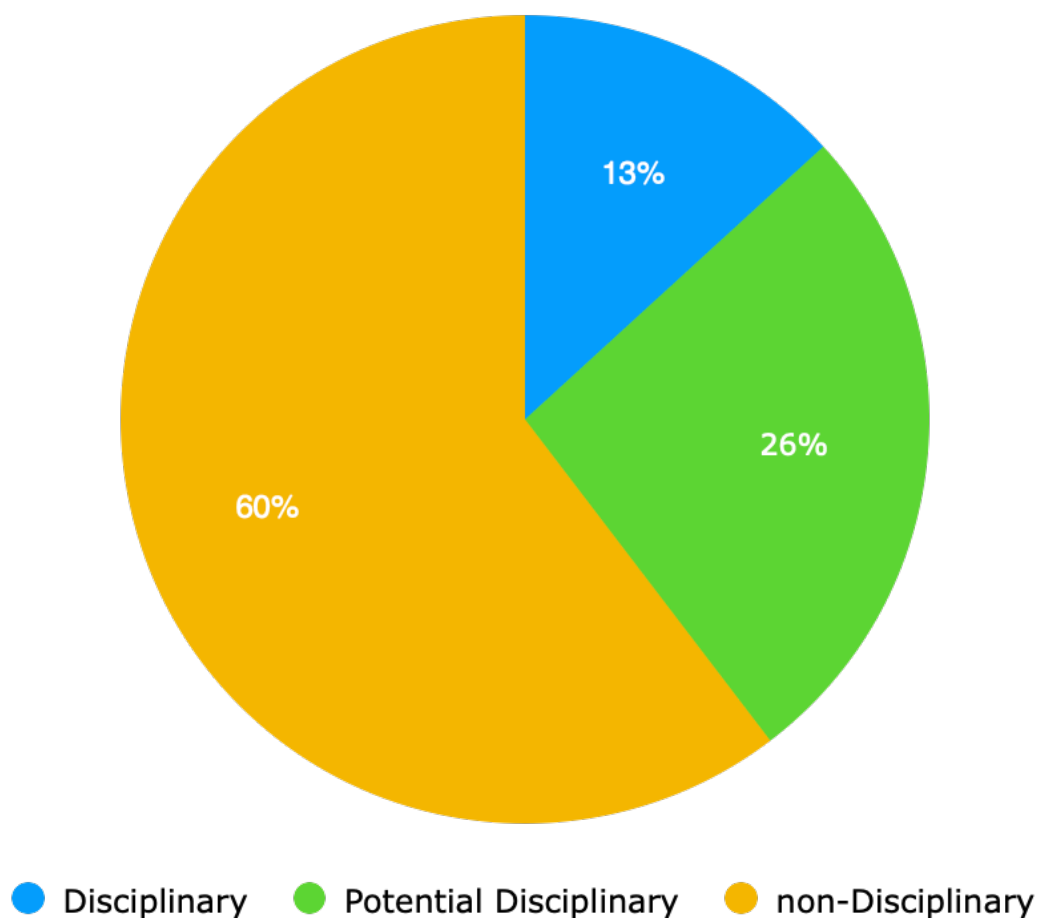
could be a reference to members of a music ensemble, such as an orchestra. These seven bundles account for a mere 1% of the total ExCo LB. Thus, there is little disciplinary content in bundles at this level. This is predictable as most high-frequency bundles consist of strings of grammatical words. Indeed, such is the case here with the 60% of LB from the most frequent 10% that are clearly not disciplinary specific (e.g., ExCo 1 'as well as').

Table 4.2.2: Disciplinary Terms in top 10% of LB

Rank	Lexical Bundle	Rank	Lexical Bundle
10	of the music	40	the music of
24	of the piece	42	members of the
29	of the song	56	the first movement
32	of the work		

The remaining group among the most frequent 10% consist of LB potentially specific to the Discourse. Accounting for 26% of this top tier, these are presented in Table 4.2.3, a perusal of which will immediately clarify why they seem to belong to the Discourse. Indicating position, order, portion or time, these bundles are all directly relevant to any discussion of a temporal art such as music. The one outlier among these is ExCo 48 'the nineteenth century', though this is likely related to the historical context of the music under discussion in the Discourse.

Table 4.2.3: Potential Disciplinary Terms

Rank	Lexical Bundle	Rank	Lexical Bundle
3	the end of	15	at the end of
5	in the first	16	the beginning of
6	part of the	28	in the second
8	end of the	30	a series of
9	of the first	39	at the beginning
11	at the end	44	half of the
13	the end of the	48	the nineteenth century

If the surmise that the top 10% of LB contain the key to significant inferences regarding the Discourse is correct, it would seem that Music Discourse: A) is indeed focused on the phenomenon of music; B) construes that phenomenon as instantiated through pieces, songs and works; C) is mostly concerned with musical compositions; D) investigates those compositions through analysis of place, order and portion of musical pieces. As the following discussion makes clear, this is an accurate understanding of this Discourse, so far as it goes. What is not evident from this cursory overview, however, is why this should be so and what the Discourse excludes; two concerns that will be addressed later.

4.3 Syntagmatic Structure

Following the above overview, LB structure is considered next through analysis of individual word classes and functions. This provides a view of the syntagmatic structure of bundles. Because they tend to straddle group and phrase boundaries, LB may be considered a form of partial syntagm. To categorize bundles by syntagmatic structure, all words are analyzed by word class and function as listed in Table 4.3.1 adapted from Halliday and Matthiessen (2014). This table includes all of the classes and functional designations for words employed in this study. Because LB often cross boundaries between functional groups, particularly

Table 4.3.1: Intersection of Word Class and Function

Class		Function				
Primary class	Secondary class	nominal group	verbal group	adverb-ial group	conjunction group	preposition group
nominal	noun	head/thing, classifier				
	pronoun	pronominal				
	adjective	post-deictic	verbal			
		epithet, classifier				
	numeral	numerative / ordinative				
		numerative / quantitative				
	determiner	deictic				
verbal	verb		verbal			
	preposition					preposition
adverb-ial	adverb			adverb-ial	conjunction	preposition
	conjunction				conjunction	

Adapted from Halliday & Matthiessen (2014:427)

Abbreviations		N	noun
Adj	adjective	Num	numeral
Adv	adverb	Prep	preposition
Conj	conjunction	Prn	pronoun
Det	determiner	V	verb

for post-qualifying prepositional phrases and verbal groups, it can be difficult to determine function; therefore, word class analysis facilitates description. To maintain uniformity of analysis, all LB are labelled by word class and then function.

A greater degree of delicacy is used to analyze nominals due to their prevalence in the Discourse and the fact that they are the primary carriers of propositional content. By contrast, high frequency verbals and adverbials found herein are wholly grammatical, containing no propositional content. Therefore, they are analyzed with a lesser degree of delicacy. General nouns have been analyzed simply as nominals regardless of their function within the nominal group, as any greater degree of delicacy would not serve to indicate a particular role within the Discourse. Nouns occurring in the terminal position of a bundle have also been analyzed this way since it is unknown in all instances whether the slot succeeding such LB contains a dissociated Thing. Additionally, due to the large number of ordinal numbers found in these LB, the word class category of Table 4.3.1 distinguishes 'numeral' as a secondary class of the primary nominal one. Using this Table as a reference, word classes for all words in the LB were analyzed and the number of each class per slot is presented in Tables 4.3.2, 4.3.3, and 4.3.4 as a

percentage of the total number of slots in all Tokens for 3-, 4-, and 5-LB.

Table 4.3.2: 5-LB Word Classes per Slot

	Slot 1		Slot 2		Slot 3		Slot 4		Slot 5	
Clas	#	%	#	%	#	%	#	%	#	%
Det	2	40%	3	60%					4	80%
N					4	80%	1	20%		
Num			2	40%	1	20%				
Prep	3	60%					4	80%	1	20%

Table 4.3.3: 4-LB Word Classes per Slot

	Slot 1		Slot 2		Slot 3		Slot 4	
Class	#	%	#	%	#	%	#	%
Adj			1	2%	3	6%		
Adv			2	4%				
Adv/ Adj							1	2%
Det	18	37%	20	41%	1	2%	21	43%
N			19	39%	16	33%	9	18%
Num	3	6%	4	8%	6	12%		
Prep	26	53%	1	2%	21	43%	14	29%
Prn			1	2%	1	2%	4	8%
V	2	4%	1	2%	1	2%		

Table 4.3.4: 3-LB Word Classes per Slot

	Slot 1		Slot 2		Slot 3	
Class	#	%	#	%	#	%
Adj	5	1%	10	2%	21	4%
Adv	24	5%	13	3%	16	3%
Adj/Adv					1	>1%
Conj	18	4%	8	2%	4	1%
Det	114	24%	104	22%	141	29%
N	71	15%	117	24%	81	17%
Adj/Adv/N			1	>1%	1	>1%
Adj/N			1	>1%	11	2%
Num	5	1%	18	4%	22	5%
Prep	142	29%	110	23%	128	27%
Adv/Prep			6	1%	4	1%
Prn	47	10%	34	7%	19	4%
V	56	12%	60	12%	33	7%

4.3.1 Structure of 5-LB

Having analyzed word classes for all LB, it is possible to distill the various Syntagmatic Types represented by these bundles (following *The Longman Grammar*, 1999). The 5-LB consist of three Syntagmatic Types, two of which account for four of the five bundles. Thus, though these common Types represent 66% of the total, or three Types of 5-LB, their occurrences account for 80% of all 5-LB. Given that all word classes in these bundles are nominal in

nature, save for prepositions, these 5-LB evidently form important nominal groups within the Discourse. The analysis of word class by slot for 5-LB is presented in Table 4.3.5, listed by corpus rank. Syntagmatic Types are listed in Table 4.3.6. Among the five Tokens in Slot 1 (S1), there are two determiners, accounting for 40% of the total, while three prepositions account for the remaining 60%. These same two word classes fill Slot 5 (S5); thus, it is the central three slots that are of interest as they each contain nouns (N) or numbers (Num). These nominals are both prevalent in the central slots (excluding determiners) and account for all five Tokens of Slot 3 (S3). An inspection of this slot reveals an intriguing pattern: the four nouns contained herein each reference a portion of something. This is analyzed in greater detail below. The relatively high percentage of ordinal numbers in these 5-LB also suggests a pattern that will become more evident with both the 4- and 3-LB.

Table 4.3.5: 5-LB Word Class Analysis

Rank	Lexical Bundle					Word Class				
	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5
78	at	the	end	of	the	Prep	Det	N	Prep	Det
218	the	se-cond	half	of	the	Det	Num	N	Prep	Det
239	at	the	begin-ning	of	the	Prep	Det	N	Prep	Det
409	the	first	half	of	the	Det	Num	N	Prep	Det
452	in	the	second	half	of	Prep	Det	Num	N	Prep

Table 4.3.6: 5-LB Syntagmatic Types (5 Tokens)

Type	#	%	Slot 1	Slot 2	Slot 3	Slot 4	Slot 5
5.1	2	40%	Prep	Det	N	Prep	Det
5.2	2	40%	Det	Num	N	Prep	Det
5.3	1	20%	Prep	Det	Num	N	Prep

4.3.2 Structure of 4-LB

Compared with the three Types of 5-LB, the 49 Tokens of the 4-LB distill into 15 Syntagmatic Types. The 4-LB word-class analysis is given in Table 4.3.7; the Syntagmatic Types are listed in Table 4.3.8. Predictably, the larger number of LB Tokens relative to the 5-LB results in a wider variety of word classes; only conjunctions are not represented here. (N.B.: Some slots contain words that may serve in more than one class, such as 'most' in ExCo 92 'one of the most'. Consequently, such instances have multiple labels.) As was true with the 5-LB, the central slots of the 4-LB are most frequently filled with nominals. In fact, 49% of all words in Slot 2 (S2) and 51% of Slot 3 (S3) are nominals. Again, the larger number of bundles generates a wider variety of nominals. In fact, the three most common Types, accounting for 56% of the 49 4-LB Tokens and 20% of all 15 Types, focus on nouns, nouns, and numeratives, respectively. While the former two are unsurprising, given the reliance of academic writing on nominalization, it is somewhat more noteworthy that numeratives are also relatively frequent since this suggests a

Table 4.3.7: 4-LB Word Class Analysis

RANK	Lexical Bundle				Word Class			
	S1	S2	S3	S4	S1	S2	S3	S4
13	the	end	of	the	Det	N	Prep	Det
15	at	the	end	of	Prep	Det	N	Prep
25	at	the	same	time	Prep	Det	Adj	N
51	on	the	other	hand	Prep	Det	Adj	N
66	in	the	case	of	Prep	Det	N	Prep
73	in	the	context	of	Prep	Det	N	Prep
74	at	the	beginning	of	Prep	Det	N	Prep
80	as	well	as	the	Prep	Adv	Prep	Det
81	the	beginning	of	the	Det	N	Prep	Det
92	one	of	the	most	Num	Prep	Det	Adv/A
150	on	the	one	hand	Prep	Det	Num	N
152	the	second	half	of	Det	Num	N	Prep
160	the	ways	in	which	Det	N	Prep	Prn
175	the	rest	of	the	Det	N	Prep	Det
207	the	way	in	which	Det	N	Prep	Prn
216	second	half	of	the	Num	N	Prep	Det
253	in	relation	to	the	Prep	N	Prep	Det
256	over	the	course	of	Prep	Det	N	Prep
260	in	the	United	States	Prep	Det	Adj	N
268	of	the	twentieth	century	Prep	Det	Num	N
286	in	a	way	that	Prep	Det	N	Prn
287	in	the	form	of	Prep	Det	N	Prep
290	the	first	half	of	Det	Num	N	Prep
291	the	first	move-ment	of	Det	Num	N	Prep
304	is	one	of	the	V	Num	Prep	Det

Table 4.3.7: 4-LB Word Class Analysis

RANK	Lexical Bundle				Word Class			
	S1	S2	S3	S4	S1	S2	S3	S4
319	as	part	of	the	Prep	N	Prep	Det
320	at	the	start	of	Prep	Det	N	Prep
323	in	the	second	half	Prep	Det	Num	N
335	the	course	of	the	Det	N	Prep	Det
344	of	the	nine-teenth	century	Prep	Det	Num	N
354	the	start	of	the	Det	N	Prep	Det
357	for	the	first	time	Prep	Det	Num	N
367	a	member	of	the	Det	N	Prep	Det
370	first	half	of	the	Num	N	Prep	Det
384	on	the	basis	of	Prep	Det	N	Prep
397	the	fact	that	the	Det	N	Prn	Det
411	the	opening	of	the	Det	N	Prep	Det
418	can	be	found	in	V	V	V	Prep
428	the	music	of	the	Det	N	Prep	Det
459	the	turn	of	the	Det	N	Prep	Det
482	as	a	means	of	Prep	Det	N	Prep
500	as	a	result	of	Prep	Det	N	Prep
510	the	context	of	the	Det	N	Prep	Det
512	the	role	of	the	Det	N	Prep	Det
521	as	well	as	a	Prep	Adv	Prep	Det
523	in	contrast	to	the	Prep	N	Prep	Det
525	in	the	nine-teenth	century	Prep	Det	Num	N
547	in	this	article	I	Prep	Prn	N	Prn
554	the	same	time	the	Det	Adj	N	Det

Table 4.3.8: 4-LB Syntagmatic Types

Type	#	%	Slot 1	Slot 2	Slot 3	Slot 4
4.1	11	22%	Det	N	Prep	Det
4.2	11	22%	Prep	Det	N	Prep
4.3	6	12%	Prep	Det	Num	N
4.4	3	6%	Det	Num	N	Prep
4.5	3	6%	Prep	N	Prep	Det
4.6	3	6%	Prep	Det	Adj	N
4.7	2	4%	Det	N	Prep	Prn
4.8	2	4%	Num	N	Prep	Det
4.9	2	4%	Prep	Adv	Prep	Det
4.10	1	2%	Det	N	Prn	Det
4.11	1	2%	Det	Adj	N	Det
4.12	1	2%	Num	Prep	Det	Adv/Adj
4.13	1	2%	Prep	Prn	N	Prn
4.14	1	2%	V	Num	Prep	Det
4.15	1	2%	V	V	V	Prep

N.B.: Rounding to the nearest decimal results in a loss of 2% of the total.

characteristic of this Discourse, one that will again be noted at the level of the 3-LB. Indeed, the third most frequent Syntagmatic Type includes a numerative in the S3 position as a pre-modifier to the nominal in S4.

4.3.3 Structure of 3-LB

Finally, Table 4.3.9 lists only the first 32 syntagmatic Types of 3-LB. The remaining 70 are not listed here because each constitutes less

than 1% of the total number of LB Tokens at the 3-word level. As with the 4-LB, the three most frequent Syntagmatic Types (S-Types) contain nouns. Though these common Types represent only 3% of the 102 S-Types of 3-LB, their occurrences account for 41% of all 3-LB. Additional S-Types from these most frequent 32 containing nouns are: 3.6, 3.10, 3.13, 3.14, 3.24, 3.27, 3.30. Numeratives rank highly again, being included in S-Type 3.5, though less frequently in S-Types 3.10 and 3.27. Of the 70 S-Types excluded from Table 4.3.9, the following eight also contain numeratives: 3.21, 3.29, 3.32, 3.34, 3.44, 3.45, 3.64, 3.95. This indicates that numeratives remain significant in the Discourse even at less frequent levels. Thus, the 3-LB exhibit patterns of noun and numerative content similar to those observed in the 4- and 5-LB. As with the 4- and 5-LB, the central slot (S2) is most productive for nominal content with 30% of all Tokens containing a central noun, though now the final slot (S3) is nearly as filled by nominals at 28%. Even 17% of S1 contains nominals. Thus, the 3-LB have a greater potential to contain disciplinary terms than either the 4- or 5-LB. Again, numeratives are well represented.

The fact that a small number of LB S-Types are highly recurrent is consistent with the idea that texts are formulaic (Sinclair's idiom principle, 1991; Ellis, N. C. 2003), both in the repetition of particular

Table 4.3.9: 3-LB Common Syntagmatic Types

Type	#	%	Slot 1	Slot 2	Slot 3
3.1	85	18%	Det	N	Prep
3.2	64	13%	N	Prep	Det
3.3	47	10%	Prep	Det	N
3.4	18	4%	Prep	Det	Adj
3.5	18	4%	Prep	Det	Num
3.6	14	3%	Prep	N	Prep
3.7	11	2%	V	Prep	Det
3.8	10	2%	Prep	Det	Adj/N
3.9	9	2%	Prn	V	Det
3.10	9	2%	Det	Num	N
3.11	8	2%	Prn	V	Adv
3.12	7	1%	Adv	Prep	Det
3.13	7	1%	Det	Adj	N
3.14	7	1%	Prep	Prn	N
3.15	7	1%	Prn	Prep	Det
3.16	7	1%	V	V	Prep
3.17	7	1%	V	V infin	V infin
3.18	6	1%	Conj	Prep	Det
3.19	6	1%	Prn	Prn	V
3.20	5	1%	Prn	V	V
3.21	5	1%	V	V	V
3.22	5	1%	V infin	V infin	Det
3.23	4	1%	Adv	Adv	Det
3.24	4	1%	Det	N	Prn
3.25	4	1%	Prep	Prn	Prn
3.26	3	1%	Conj	Prn	V
3.27	3	1%	Num	N	Prep

Table 4.3.9: 3-LB Common Syntagmatic Types

Type	#	%	Slot 1	Slot 2	Slot 3
3.28	3	1%	Prep	Prep	Det
3.29	3	1%	Prep	Prn	Det
3.30	3	1%	Prep	Adj	N
3.31	3	1%	V	V	Adv/Prep
3.32	3	1%	V	Prn	Det

LB and even syntagmatic structure. Moreover, the substantial presence of numeratives is an indication of a particular configuration of resources (i.e., nominals and numeratives) in the Discourse. As Thompson (2014) states, such configurations constitute register. Applying this idea with greater delicacy, the highly frequent interaction of nominals and numeratives in Music Discourse suggests a distinguishing feature that marks it as a sub-register of Academic Discourse. In other words, if register is a constellation of features that construe content (Halliday 2001), then a constellation of more disciplinary-specific features construes content in a manner that is still more particular than Academic Discourse. Given that all forms of academic writing have their own conventions (Johns 2003), the interaction of nominals and numeratives is one such example in Music Discourse. This will become increasingly evident at each subsequent stage of analysis, reinforcing Wray's (2012) statement that LB differ partly by discipline.

Before proceeding, a later finding relating to numeratives must be introduced here. As has already been suggested by the preceding analysis, and as will become evident later in this chapter, numeratives form a substantial and therefore significant part of the Discourse. The reasons for this will be discussed later; for now, it is sufficient to note that this is to be expected given that music is a temporal art. That is, numbers and numbering, portions and ordering, are indispensable elements in any discourse related to any process that unfolds structure over time, in this case repeatedly with each hearing. Consequently, this Discourse not only employs ordinal numbers but also cardinal numbers and generic nouns that measure quantity. Such items that function as numeratives can be either definite or indefinite, and can quantify or order, as shown by the examples in Table 4.3.10 from Halliday and Matthiessen (2014:375).

Table 4.3.10: Items functioning as Numerative

	Definite	Indefinite
quantitative	one, two, three, etc. [a couple of], etc. [a quarter of], etc.	few, little, [a bit of], etc. several, [a number of], etc.
ordinative	first, second, third, etc. next, last	preceding, subsequent, etc.

Adapted from Halliday and Matthiessen (2014: p.375)

4.4 Word Functions

Analyzing the word class of each LB slot enabled distillation of Syntagmatic Types, which revealed that the most frequent S-Types contain nouns and numbers. The former word class aligns with the fact that academic writing relies heavily on nominals and nominalization to communicate ideas. The latter class suggests that sequential ordering is an integral process in Music Discourse. Taken together, these two high-frequency word classes hint at a particular configuration of lexical features in the Discourse. To further investigate the suggestion that ordering is integral to it, an analysis of LB functionality is presented next by 5-, 4-, and finally 3-LB.

Note that the only functional group to be analyzed to the greatest degree of delicacy is the nominal one, as this is most relevant to exploring the disciplinary content of LB. (N.B.: Because it was impossible in many instances to discern whether an LB noun functioned as Head or Thing, given the lack of context, such nouns are simply analyzed as 'Nominal'.)

4.4.1 5-LB Word Functions

The functional analysis of numeratives offers a greatly expanded view of them from that presented by word class analysis alone. From this enlarged perspective, a variety of word classes can

function as numeratives. This is significant for Music Discourse because it is replete with such items, a fact already intimated in the functional analysis of the five Tokens of the 5-LB, presented in Table 4.3.11. Indeed, each of these five bundles contains one or more words functioning as a numerative: two contain definite ordinatives (ExCo 78 'end', ExCo 239 'beginning'), and three contain complex numeratives composed of a definite ordinative followed by the definite quantitative 'half' (ExCo 218, 409, 452). The fact that all Tokens of the 5-LB contain numeratives strongly suggests that both sequential ordering and quantifying are fundamental procedures in the Discourse. This inference is especially supported by the rank of these bundles among the total LB: the higher frequency bundles are those enumerating the latter portions of something, while the lower frequency ones enumerate former portions, which is indicative of sequencing. In other words, ordering is a fundamental process of the Discourse. Indeed, the presence of 'half' in three of the five Tokens shows a concern with form. Table 4.3.12 displays the slots containing numeratives in the 5-LB, showing that 32% of all 25 slots are filled with words functioning thus.

Table 4.3.11: 5-LB Word Function Analysis

Rank	Lexical Bundle					Function				
	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5
78	at	the	end	of	the	pre-posi-tion	deic-tic	ordi-na-tive	pre-posi-tion	deic-tic
218	the	se-cond	half	of	the	deic-tic	compound numerative (ordinative, quantita-tive)		pre-posi-tion	deic-tic
239	at	the	begin-ning	of	the	pre-posi-tion	deic-tic	ordi-na-tive	pre-posi-tion	deic-tic
409	the	first	half	of	the	deic-tic	compound numerative (ordinative, quantita-tive)		pre-posi-tion	deic-tic
452	in	the	se-cond	half	of	pre-posi-tion	deic-tic	compound numerative (ordinative, quantita-tive)		pre-posi-tion

Table 4.3.12: 5-LB Numeratives by Slot

Function	% of Total	Slot 2	Slot 3	Slot 4
complex numerative	60%	2		
				1
ordinative	40%		2	

4.4.2 4-LB Word Functions

As with the word-class analysis, the accuracy of the above inference based on 5-LB can be tested by analyzing the functionality of 4-LB. Table 4.3.13 gives the complete word-function analysis for all 49 Tokens of 4-LB. Of these, no fewer than 22 contain numeratives, which accounts for 45% of all 4-LB. Such a high percentage (nearly half) bolsters the view that the Discourse is significantly dependent on numeratives. These 22 bundles are divided into compound numeratives, ordinatives and quantitatives, as shown in Table 4.3.14. The ordinative group contains the largest percentage of high-frequency bundles, the first four of which replicate the precedent observed among the 5-LB of 'end' (first two bundles) before 'beginning' (next two); 'start' and 'opening' also appear as lower frequency bundles in this group. The same pattern repeats among the compound numeratives, with 'second half' taking precedence over 'first half'. While this is clearly to be expected in any discipline concerned with a temporal art, given the need to describe how such art unfolds over time, it may be less obvious that this could also be an indication of the Discourse's dependence on staff notation, as scores can be conveniently analyzed using temporal language to reference spatial locations on the page.

Table 4.3.13: 4-LB Word Functions

Rank	Lexical Bundle				Function			
	S1	S2	S3	S4	S1	S2	S3	S4
13	the	end	of	the	deic-tic	ordi-na-tive	pre-posi-tion	deic-tic
15	at	the	end	of	pre-posi-tion	deic-tic	ordi-na-tive	pre-posi-tion
25	at	the	same	time	pre-posi-tion	deic-tic	post-deic-tic	nomi-nal
51	on	the	other	hand	pre-posi-tion	deic-tic	post-deic-tic	nomi-nal
66	in	the	case	of	pre-posi-tion	deic-tic	nomi-nal	pre-posi-tion
73	in	the	con-text	of	pre-posi-tion	deic-tic	nomi-nal	pre-posi-tion
74	at	the	begin-ning	of	pre-posi-tion	deic-tic	ordi-na-tive	pre-posi-tion
80	as	well	as	the	complex preposition			deic-tic
81	the	begin-ning	of	the	deic-tic	ordi-na-tive	pre-posi-tion	deic-tic
92	one	of	the	most	quan-tita-tive	pre-posi-tion	deic-tic	ad-verb-ial/quan-tita-tive

Table 4.3.13: 4-LB Word Functions

Rank	Lexical Bundle				Function			
	S1	S2	S3	S4	S1	S2	S3	S4
150	on	the	one	hand	preposition	deictic	ordinative	nominal
152	the	second	half	of	deictic	compound numerative (ordinative, quantitative)		preposition
160	the	ways	in	which	deictic	nominal	preposition	deictic
175	the	rest	of	the	deictic	quantitative	preposition	deictic
207	the	way	in	which	deictic	nominal	preposition	deictic
216	second	half	of	the	compound numerative (ordinative, quantitative)		preposition	deictic
253	in	relation	to	the	complex conjunction			deictic
256	over	the	course	of	preposition	deictic	nominal	preposition
260	in	the	United	States	preposition	deictic	compound thing	
268	of	the	twentieth	century	preposition	deictic	compound numerative (ordinative, quantitative)	

Table 4.3.13: 4-LB Word Functions

Rank	Lexical Bundle				Function			
	S1	S2	S3	S4	S1	S2	S3	S4
286	in	a	way	that	preposition	deictic	nominal	deictic
287	in	the	form	of	preposition	deictic	nominal	preposition
290	the	first	half	of	deictic	compound numerative (ordinative, quantitative)		preposition
291	the	first	movement	of	deictic	compound numerative (ordinative, quantitative)		preposition
304	is	one	of	the	verbal	ordinative	preposition	deictic
319	as	part	of	the	preposition	quantitative	preposition	deictic
320	at	the	start	of	preposition	deictic	ordinative	preposition
323	in	the	second	half	preposition	deictic	compound numerative (ordinative, quantitative)	
335	the	course	of	the	deictic	nominal	preposition	deictic
344	of	the	nineteenth	century	preposition	deictic	compound numerative (ordinative, quantitative)	

Table 4.3.13: 4-LB Word Functions

Rank	Lexical Bundle				Function			
	S1	S2	S3	S4	S1	S2	S3	S4
354	the	start	of	the	deic-tic	ordi-na-tive	pre-posi-tion	deic-tic
357	for	the	first	time	pre-posi-tion	deic-tic	ordi-na-tive	nomi-nal
367	a	mem-ber	of	the	deic-tic	nomi-nal	pre-posi-tion	deic-tic
370	first	half	of	the	compound numerative (ordinative, quantitative)	pre-posi-tion	deic-tic	
384	on	the	basis	of	pre-posi-tion	deic-tic	nomi-nal	pre-posi-tion
397	the	fact	that	the	deic-tic	nomi-nal	deic-tic	deic-tic
411	the	open-ing	of	the	deic-tic	ordi-na-tive	pre-posi-tion	deic-tic
418	can	be	found	in	verbal group			pre-posi-tion
428	the	music	of	the	deic-tic	nomi-nal	pre-posi-tion	deic-tic
459	the	turn	of	the	deic-tic	nomi-nal	pre-posi-tion	deic-tic
482	as	a	means	of	pre-posi-tion	deic-tic	nomi-nal	pre-posi-tion

Table 4.3.13: 4-LB Word Functions

Rank	Lexical Bundle				Function			
	S1	S2	S3	S4	S1	S2	S3	S4
500	as	a	result	of	ad-verb-ial	deic-tic	nomi-nal	pre-posi-tion
510	the	con-text	of	the	deic-tic	nomi-nal	pre-posi-tion	deic-tic
512	the	role	of	the	deic-tic	nomi-nal	pre-posi-tion	deic-tic
521	as	well	as	a	complex preposition			deic-tic
523	in	con-trast	to	the	complex preposition			deic-tic
525	in	the	nine-teenth	century	pre-posi-tion	deic-tic	compound numerative (ordinative, quantitative)	
547	in	this	article	I	pre-posi-tion	deic-tic	thing	pro-nomi-nal
554	the	same	time	the	deic-tic	pre-posi-tion	nomi-nal	deic-tic

Of particular note among the compound numeratives is the presence of three references to centuries (ExCo 268 'twentieth century', 344 and 525 'nineteenth century') and one to a musical movement (ExCo 291 'first movement'). While the latter offers more concrete evidence that numeratives are vital to Music Discourse and

Table 4.3.14: 4-LB functioning as Numeratives

4-LB functioning as Numeratives					
Function	Rank	S1	S2	S3	S4
com-pound numer-ative	152	the	second	half	of
	216	second	half	of	the
	268	of	the	twentieth	century
	290	the	first	half	of
	291	the	first	movement	of
	323	in	the	second	half
	344	of	the	nineteenth	century
	370	first	half	of	the
	525	in	the	nineteenth	century
ordi-native	13	the	end	of	the
	15	at	the	end	of
	74	at	the	beginning	of
	81	the	beginning	of	the
	150	on	the	one	hand
	304	is	one	of	the
	320	at	the	start	of
	354	the	start	of	the
	357	for	the	first	time
	411	the	opening	of	the
quanti-tative	92	one	of	the	most
	175	the	rest	of	the
	319	as	part	of	the

likely employed for discussion of scores, the former seem to

contradict that assertion since these enumerate historical rather

than musical time. This then adds depth to the perspective of the primary concerns within the Discourse; at some level, it is historically oriented. This will come as no surprise to musicians and Musicologists accustomed to this Discourse.

Here, it is necessary to pause and note one problem with the analysis of these 4-LB numeratives, a problem that will recur in the analysis of the 3-LB. ExCo 150 'on the one hand' contains the ordinative 'one' yet does not function as a numerative in the same sense as the other 4-LB numeratives; rather, it is an instance of a discourse marker. This will be discussed in Section 4.9 through the bundle function analysis that employs Biber et al.'s (2004) taxonomy of LB discourse functions. As will also be seen later, other 3-LB containing the word 'hand' do function as numeratives since they reference parts of a piano score.

Having surveyed the 22 numeratives among all 49 4-LB, a few other items of interest may be noted here. Among the remaining 27 LB, there are three complex prepositions (ExCo 80, 521, 523), the first of which is the high-frequency bundle 'as well as the'. As will be seen later, this same complex preposition is an exceptionally high-frequency 3-LB. Among these remaining bundles there is also one compound Thing (ExCo 260 'in the United States'), and one Thing

and one pronominal together (ExCo 547 'in this article I'). Though half of the journal articles in the ExCo are American publications, the reference to the United States still comes as a surprise given that the Discourse is by no means focused exclusively on music of a single country, drawing as it does on composers from several countries, and given that this Corpus includes a selection of articles from Ethnomusicology journals. While it is possible that this curiosity may point to a defect in the balance of the Expert Corpus, a more likely explanation is that the selection of Music Theory articles caused this particular LB to rise to a high rank. This is because many of the articles in those journals are focused on relatively new music from the twentieth century to the present, and much of that compositional activity has taken place in, or been connected to, music in a handful of American cosmopolitan centers, most notably New York. Had more Theory journals been included in the Corpus, the presence of this bundle may well have been magnified. As will be seen, this explanation is supported by the presence of 'United States' and 'New York' among the 3-LB. (N.B.: This may also point to a sense of prestige attached to music in the U.S. and particularly New York on account of the economic significance of both globally.) Finally, regarding the last bundle mentioned above, ExCo 547 is unique in the Corpus both for being

self-referential ('this article') and containing a first-person singular pronoun.

Table 4.3.15 presents nominal content by slot in the 4-LB. The highest percentage of such content by slot consists of nominals functioning as either Head or Thing in the nominal group. Following these, numeratives account for nearly the same percentage of such content by slot. In both cases, and as is true in the 5-LB, the central slots are the most productive for nominals. The reasons for this will be discussed in the following section on prepositions and (potential) genitive bundles.

Table 4.3.15: 4-LB Nominals by Slot

Function	% of Slots	Slot 1	Slot 2	Slot 3	Slot 4
nominal	45%		9	9	4
compound thing	2%			1	
thing	2%			1	
pronominal	2%				1
complex numerative	18%	2			
			3		
				4	
ordinative	20%		5	5	
quantitative	6%	1	2		

4.4.3 3-LB Word Functions

As with the word-class analysis, the functional analysis of the 3-LB is too extensive to list here in tabular format. Nevertheless, evidence supporting the significance of numeratives in the Discourse can be presented here, as well as other items of interest. Among the latter, several 3-LB contain or consist of compound and complex groupings. As Halliday and Matthiessen (2014:423) note, with such items (e.g., multi-word prepositions), it is unnecessary to analyze them in any greater detail; a single functional designation is sufficient. For the purposes of this study, I have distinguished complex from compound items based on word class; complexes are formed by words from different classes, while compounds are formed by words of the same class, which may or may not be paratactically linked by a conjunction. These items, then, include: five complex adverbials (e.g., ExCo 318 'as much as'), fourteen complex conjunctions (e.g., ExCo 53 'in addition to'), fourteen complex prepositions (e.g., ExCo 524 'in front of'), seven compound adverbials (e.g., ExCo 192 'more or less'), and compound prepositions (e.g., ExCo 62 'back to'). Among the complex prepositions, ExCo 1 'as well as' is a bundle subsumed under the previously encountered 4-LB ExCo 80 'as well as a'. As the highest frequency bundle in the Corpus, it is clearly considered highly useful, and likely is in all forms of academic discourse.

Other items typical of academic discourse include the three discourse bundles also related to a 4-LB counterpart (i.e., ExCo 51 'on the other hand'): ExCo 52 'the other hand', 139 'on the one', and 151 'the one hand'. Then there are also eight self-referential bundles — including ExCo 97 'in this article', 112 'shown in example', 230 'of this article' — among which are those that include the author: ExCo 212 'I do not', 421 'I argue that', 415 'as we have', 430 'this article I', 494 'we have seen'. Interestingly, both 415 and 494 enlist readers in the author's observations. This may be of particular interest to novices if they have been discouraged from using first-person pronouns, specifically the plural, which could inadvertently suggest multiple writers.

After the first-person pronoun LB listed above, there are an additional five LB containing pronouns: ExCo 228 'in which he', 405 'of his own', 440 'that he had', 473 'that he was', 533 'that they were'. These will be analyzed later as a group. Having considered all of the above 3-LB, two large groups remain to be analyzed: bundles containing numeratives and disciplinary terminology.

Table 4.3.16: 3-LB containing Numeratives

Rank	Lexical Bundle			Function		
	Slot 1	Slot 2	Slot 3	Slot 1	Slot 2	Slot 3
2	one	of	the	quanti-tative	preposi-tion	deictic
3	the	end	of	deictic	ordinative	preposi-tion
5	in	the	first	preposi-tion	deictic	ordinative
6	part	of	the	quanti-tative	preposi-tion	deictic
8	end	of	the	ordina-tive	preposi-tion	deictic
9	of	the	first	preposi-tion	deictic	ordinative
11	at	the	end	preposi-tion	deictic	ordinative
16	the	begin-ning	of	deictic	ordinative	preposi-tion
17	some	of	the	quanti-tative	preposi-tion	deictic
28	in	the	second	preposi-tion	deictic	ordinative
39	at	the	begin-ning	preposi-tion	deictic	ordinative
44	half	of	the	quanti-tative	preposi-tion	deictic
46	as	part	of	preposi-tion	quanti-tative	preposi-tion
48	the	nine-teenth	century	deictic	complex numerative (ordinative, quantitative)	

Table 4.3.16: 3-LB containing Numeratives

Rank	Lexical Bundle			Function		
	Slot 1	Slot 2	Slot 3	Slot 1	Slot 2	Slot 3
56	the	first	move-ment	deictic	complex numerative (ordinative, quantitative)	
57	many	of	the	quanti-tative	preposi-tion	deictic
61	of	the	second	preposi-tion	deictic	ordinative
70	begin-ning	of	the	ordina-tive	preposi-tion	deictic
84	in	the	early	preposi-tion	deictic	quanti-tative
88	be-tween	the	two	preposi-tion	deictic	quanti-tative
93	the	twen-tieth	century	deictic	complex numerative (ordinative, quantitative)	
106	of	the	two	preposi-tion	deictic	quanti-tative
107	most	of	the	quanti-tative	preposi-tion	deictic
108	the	first	two	deictic	complex numerative (ordinative, quantitative)	
109	the	second	half	deictic	complex numerative (ordinative, quantitative)	
110	is	one	of	verbal	ordinative	preposi-tion
111	part	of	a	quanti-tative	preposi-tion	deictic
114	the	start	of	deictic	quanti-tative	preposi-tion

Table 4.3.16: 3-LB containing Numeratives

Rank	Lexical Bundle			Function		
	Slot 1	Slot 2	Slot 3	Slot 1	Slot 2	Slot 3
115	the	rest	of	deictic	quanti-tative	preposi-tion
119	of	the	open-ing	preposi-tion	deictic	ordinative
121	all	of	the	quanti-tative	preposi-tion	deictic
126	the	opening	of	deictic	ordinative	preposi-tion
128	for	the	first	preposi-tion	deictic	ordinative
145	second	half	of	complex numerative (ordinative, quantitative)		preposi-tion
172	much	of	the	quanti-tative	preposi-tion	deictic
173	rest	of	the	quanti-tative	preposi-tion	deictic
180	in	the	middle	preposi-tion	deictic	ordinative
188	the	first	half	deictic	complex numerative (ordinative, quantitative)	
201	the	first	time	deictic	ordinative	thing
204	the	first	and	deictic	ordinative	conjunc-tion
215	of	the	three	preposi-tion	deictic	quantita-tive
245	of	the	twen-tieth	preposi-tion	deictic	ordinative

Table 4.3.16: 3-LB containing Numeratives

Rank	Lexical Bundle			Function		
	Slot 1	Slot 2	Slot 3	Slot 1	Slot 2	Slot 3
252	early	twen-tieth	century	complex numerative (ordinative, ordinative, quantitative)		
254	in	the	third	preposi-tion	deictic	ordinative
258	first	half	of	complex numerative (ordinative, quantitative)		preposi-tion
266	as	one	of	preposi-tion	ordinative	preposi-tion
270	section	of	the	quanti-tative	preposi-tion	deictic
273	at	the	start	preposi-tion	deictic	ordinative
276	open-ing	of	the	ordina-tive	preposi-tion	deictic
285	first	move-ment	of	complex numerative (ordinative, quantitative)		preposi-tion
302	in	the	open-ing	preposi-tion	deictic	classifier/ ordinative
306	of	the	nine-teenth	preposi-tion	deictic	ordinative
307	the	first	of	deictic	ordinative	preposi-tion
313	start	of	the	ordina-tive	preposi-tion	deictic
338	all	of	these	quanti-tative	preposi-tion	deictic
351	of	the	century	preposi-tion	deictic	quantita-tive

Table 4.3.16: 3-LB containing Numeratives

Rank	Lexical Bundle			Function		
	Slot 1	Slot 2	Slot 3	Slot 1	Slot 2	Slot 3
363	of	the	final	preposition	deictic	classifier/ ordinative
364	the	eigh-teenth	century	deictic	complex numerative (ordinative, quantitative)	
366	to	the	first	preposition	deictic	ordinative
375	the	middle	of	deictic	ordinative	preposition
377	and	the	second	conjunction	deictic	ordinative
392	from	the	first	preposition	deictic	ordinative
398	the	first	part	deictic	complex numerative (ordinative, quantitative)	
435	in	the	nine-teenth	preposition	deictic	ordinative
448	first	and	second	compound ordinative		
454	late	nine-teenth	century	complex numerative (ordinative, ordinative, quantitative)		
480	with	the	first	preposition	deictic	ordinative
481	and	the	first	conjunction	deictic	ordinative
489	parts	of	the	quantitative	preposition	deictic
505	of	the	prev-ious	preposition	deictic	classifier/ ordinative

Table 4.3.16: 3-LB containing Numeratives

Rank	Lexical Bundle			Function		
	Slot 1	Slot 2	Slot 3	Slot 1	Slot 2	Slot 3
519	all	the	more	quanti-tative	deictic	adverbial
520	any	of	the	quanti-tative	preposi-tion	deictic
530	on	the	first	preposi-tion	deictic	ordinative
535	the	second	and	deictic	ordinative	conjunc-tion
549	of	the	four	preposi-tion	deictic	quanti-tative

As with the 5- and 4-LB, numeratives account for a significant portion of the 482 3-LB, as shown in Table 4.3.16. Altogether, there are a total of 75 numeratives, or nearly 16% of all 3-LB, again confirming findings from the 5- and 4-LB suggesting that numeratives are a central feature of the Discourse. The distribution of these 75 LB Types across various numerative functions is shown in Chart 4.3.1. The function of these bundles in the Discourse is analyzed in greater detail later. Here, it is sufficient to note that order and quantity are the two primary functions of numeratives in these 3-LB. Complex numeratives refer to LB containing more than one type of numerative, such as an ordinative followed by a quantitative (e.g., ExCo 145 'second half of'). Though both words

share a function, these items are analyzed as complex rather than compound to highlight the fact that each numerative has a different function within the numerative category.

Chart 4.3.1: 3-LB containing Numeratives
(75 Tokens)

- Ordinative
- Quantitative
- Complex Numerative
- Compound Ordinative
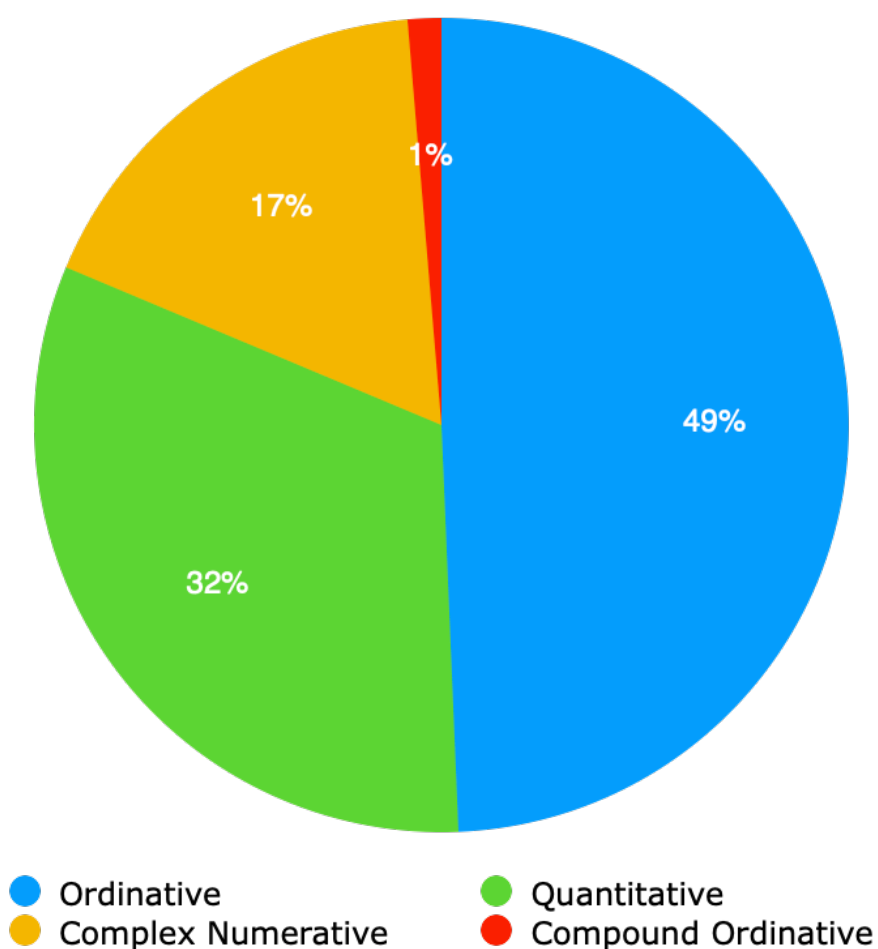
Table 4.3.17 presents nominal word functions by slot in 3-LB. While a more detailed analysis of this nominal content is presented below, it is noteworthy here that other nominals considerably outweigh numeratives. This raises the possibility that this group of LB may contain a substantial amount of disciplinary content. In total, 24%

Table 4.3.17: 3-LB Nominal Word Functions by Slot

Function	% of Slots	Slot 1	Slot 2	Slot 3
classifier	2%		2	6
classifier/ nominal	>1%			1
classifier/ ordinative	1%			5
classifier/ thing	>1%			2
epithet	2%	1	1	8
epithet/ nominal	>1%			1
epithet/thing	>1%			1
nominal	32%	48	80	27
compound thing	1%		3	
thing	13%	10	10	42
thing/ classifier	>1%			1
pronominal	1%	3	4	
complex numerative 3-slots	>1%		2	
complex numerative 2-slots	1%	3		
complex numerative 2-slots	2%		8	
compound ordinative	>1%		1	
ordinative	8%	4	12	21
quantitative	4%	15	6	

(348) of all 1,446 slots of these bundles are filled with nominals.

Nonetheless, numeratives still account for a significant portion of the total slots at 17%, which is the second largest, again confirming their importance in the Discourse.

Lastly, words functioning as verbs and verbal groups are excluded from this discussion because, with two (potential) exceptions, they are not disciplinary-specific: ExCo 155 'to create a', potentially referencing a composition or specific feature thereof; ExCo 295 'be heard as'. Aside from these, all other verbs and verbal groups consist solely of auxiliaries, conditionals, or forms of 'to be'. Thus, they cannot be distinguished as specifically musical rather than generally academic.

4.4.4 Prepositions and (Potential) Genitive LB

Until now, word classes and functions have been considered with a focus both on general nominals and the frequency of numeratives in the Discourse. The purpose of this discussion has been to highlight potential disciplinary content (i.e., nominals) and a salient discoursal feature (i.e., numeratives). In this section, prepositions within all of the LB (3-, 4-, 5-LB) are discussed by frequency, after which bundles that are genitive or potentially genitive are examined. As a closed word class, prepositions are naturally high-frequency items in any discourse. As the head of a prepositional phrase, particularly

rank-shifted post-qualifying ones in nominal groups, they saturate academic writing, partly on account of its reliance on nominalization. Prepositions function as minor processes; that is, they properly relate more closely to the verbal group than any other functional group (Halliday & Matthiessen 2014:423).

There is one preposition within this word class that can have a unique function: 'of'. As Halliday and Matthiessen (ibid.:392) note, this preposition can indicate a structural relationship between nominals, allowing the functional Head of the nominal group to be dissociated from the Thing. Where this occurs, the Head of the group may be a generic noun while the dissociated Thing offers specific semantic content. Examples of this from the Expert Corpus are most readily found in the ambient content attached to the 5-LB; for instance, the second Token of ExCo 78 ('at the end of the'): 'at the end of the *song'*. In this instance, the Thing is 'song', not 'end'. This dissociation of Head and Thing allows the Thinghood of 'song' to be mapped onto the Headhood of 'end' (ibid.:394), thus extending the degree of specificity possible within the nominal group by adding a further grammatical resource for relating nominals to one another.

In the ExCo, 'of' can be found in every slot, from one to five. Thus, it is an especially productive grammatical item. This can be demonstrated by the frequency at which it appears in the Corpus. Out of 442 prepositions in ExCo LB (one per bundle for ⅘ of the total Types), 'of' accounts for half of these, as shown in Chart 4.4.1. The next highest frequency preposition is 'in' (96 LB) at 22%. The remaining prepositions account for far smaller percentages of the total 442 bundles that contain prepositions. The high percentage of LB containing 'of' indicate that numerous ExCo LB have the potential to form the type of dissociative structural relationship outlined above or to function as genitives without such dissociation. The former potential opens the possibility that there could be a substantial number of extended numeratives either within these LB or in concert with their ambient content. Therefore, these bundles are discussed below in order by slot, from initial, to medial, to terminal slots.

Table 4.4.1 displays the ten most frequent LB Types with 'of' in the initial slot. These are categorized into three groups: disciplinary, numerative, and other. The complete list of 42 initial slot 'of' are not presented here in order to save space; however, the percentage of the total 42 Types represented by each category is listed in Chart 4.4.2. While the category of 'other' is largest, the remaining two

(i.e., disciplinary and numerative) account for nearly two-thirds of the total. By viewing only these bundles, it is unknown what precedes them; thus, it is uncertain whether these all consistently form dissociated Head-Thing structures. Therefore, these have the potential to function as genitives.

Chart 4.4.1: Prepositions (% of 442 Types)
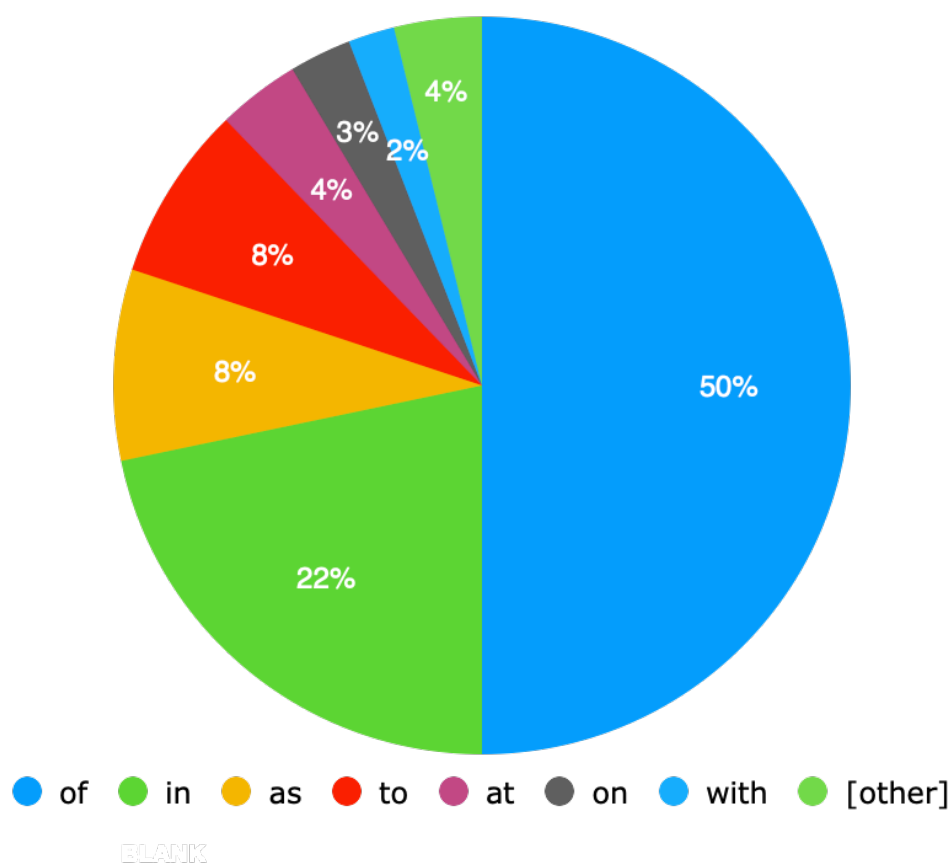
Table 4.4.2 presents 62 of the total 84 Types of bundles with 'of' in a medial slot. These 62 have the potential to form extended numeratives, whereas the 22 that have been excluded from analysis do not. As 3-, 4- and 5-LB are presented together here, medial slots

Table 4.4.1: Initial Slot 'of' (Top 10 Types)

Rank	Slot 1	Slot 2	Slot 3	Type
9	of	the	first	numerative
10	of	the	music	disciplinary
24	of	the	piece	disciplinary
29	of	the	song	disciplinary
32	of	the	work	disciplinary
36	of	the	most	numerative
61	of	the	second	numerative
64	of	the	musical	disciplinary
106	of	the	two	numerative
117	of	the	same	other

Chart 4.4.2: Initial Slot 'of' Functions
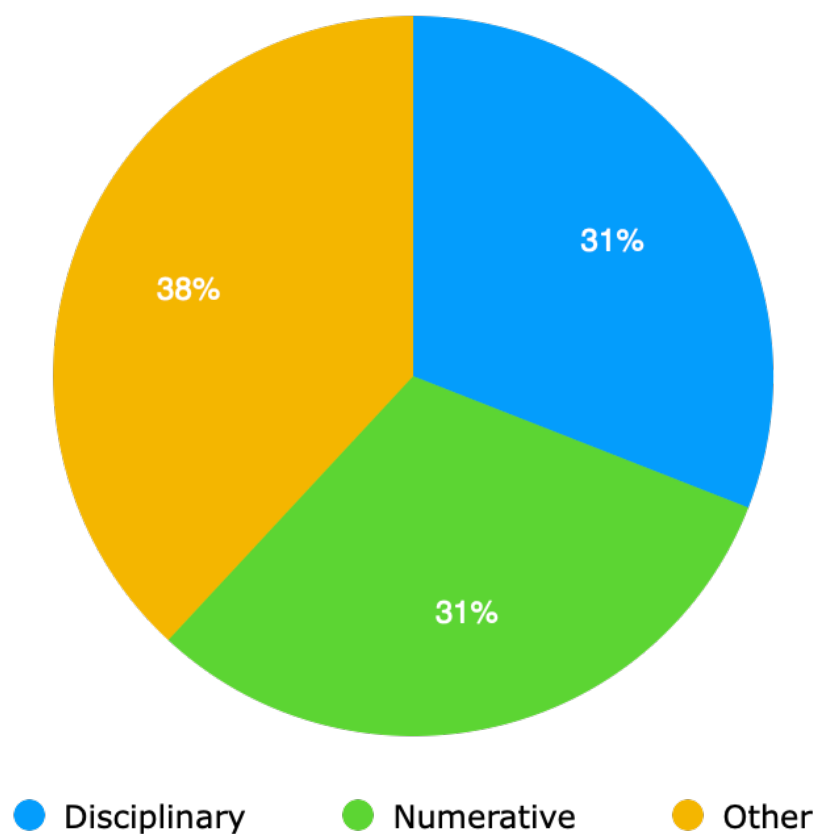
31% Disciplinary
31% Numerative
38% Other

Table 4.4.2: Medial Slot 'of' (84 Types)

Rank	Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Type
2	one	of	the			quantum
6	part	of	the			portion
8	end	of	the			portion
13	the	end	of	the		portion
17	some	of	the			portion
42	members	of	the			quantum
44	half	of	the			portion
57	many	of	the			portion
70	beginning	of	the			portion
78	at	the	end	of	the	portion
81	the	begin-ning	of	the		portion
92	one	of	the	most		quantum
103	version	of	the			variety
105	music	of	the			facet
107	most	of	the			portion
111	part	of	a			portion
121	all	of	the			portion
132	each	of	the			quantum
136	nature	of	the			make-up
143	aspects	of	the			facet
168	member	of	the			quantum
172	much	of	the			portion
173	rest	of	the			portion
175	the	rest	of	the		portion
182	structure	of	the			make-up
216	second	half	of	the		portion

Table 4.4.2: Medial Slot 'of' (84 Types)

Rank	Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Type
218	the	second	half	of	the	portion
221	movement	of	the			quantum
239	at	the	begin-ning	of	the	portion
246	sense	of	the			make-up
250	under-standing	of	the			make-up
264	analysis	of	the			disciplinary
267	course	of	the			facet
270	section	of	the			portion
276	opening	of	the			portion
304	is	one	of	the		quantum
313	start	of	the			portion
319	as	part	of	the		portion
335	the	course	of	the		facet
338	all	of	these			portion
354	the	start	of	the		portion
355	versions	of	the			variety
367	a	mem-ber	of	the		quantum
370	first	half	of	the		portion
381	interpreta-tion	of	the			variety
388	those	of	the			portion
390	voice	of	the			facet
393	idea	of	the			make-up
396	perform-ance	of	the			quantum

Table 4.4.2: Medial Slot 'of' (84 Types)

Rank	Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Type
409	the	first	half	of	the	portion
411	the	opening	of	the		portion
416	aspect	of	the			facet
419	develop-ment	of	the			facet
428	the	music	of	the		facet
432	each	of	these			portion
439	sound	of	the			facet
465	case	of	the			variety
470	level	of	the			make-up
489	parts	of	the			portion
507	quality	of	the			variety
520	any	of	the			portion
544	content	of	the			make-up

may be in any of three potential medial positions, depending on the number of words in a given bundle. Chart 4.4.3 shows the percentage of each group among these LB. In contrast to bundles with initial 'of', these bundles can be grouped more precisely as the content of the slot preceding 'of' is present, making it easier to determine whether each Type represents a dissociated structural relationship or not. Hence, extended numeratives are analyzed herein with a greater degree of delicacy. Bundles not containing

numeratives or disciplinary content have been excluded from Table 4.4.3 (e.g., ExCo 83 'because of the').

Chart 4.4.3: Medial Slot 'of' Functions
(% of 62 Types)
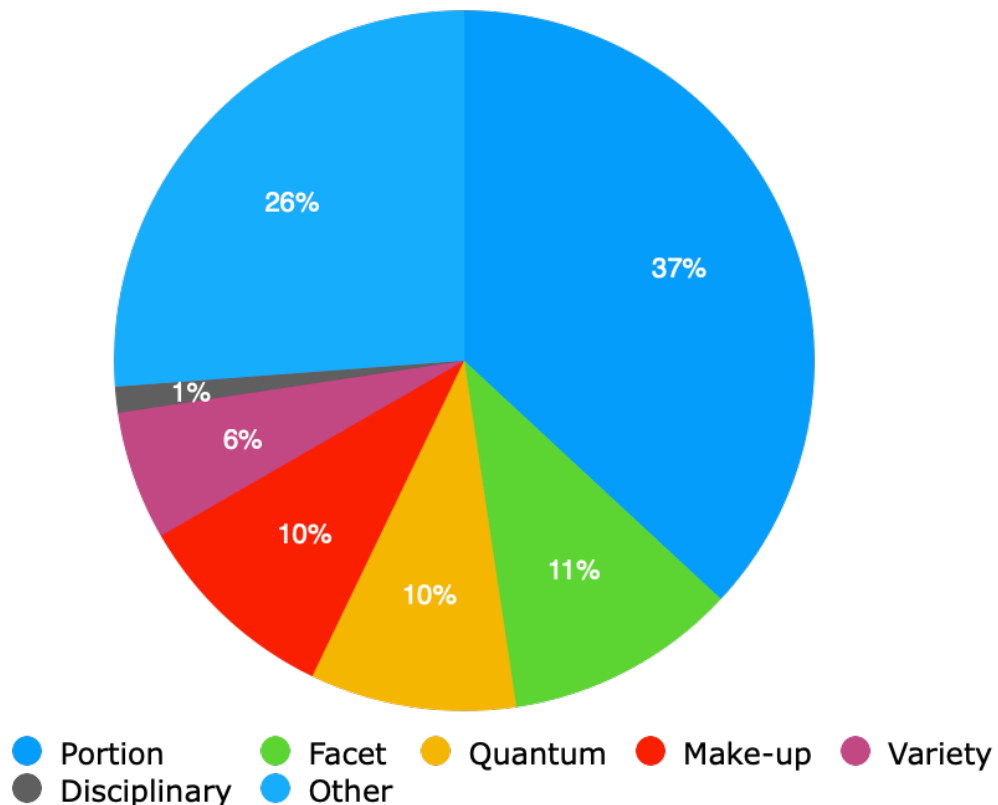
These bundles with medial 'of' are more productive than those with initial 'of' largely because they include 4- and 5-word bundles. Additionally, several of the initial nominals are dissociated Heads; without their dissociated Thing, it is not possible to generalize their use as part of disciplinary-specific terminology. Moreover, extended numeratives comprise the bulk of these bundles (76%), with the majority being quantitative measures of portion. Interestingly, there

are no occurrences of aggregate numeratives here, suggesting either a top-down view of musical pieces in the Discourse or a focus on macro-structural strata just below the level of a complete piece. Since the portion bundles likely all focus on score analysis, it is reasonable to infer that Music Discourse prioritizes consideration of sections of works over complete pieces. This may be in service of a larger goal to describe or analyze an entire work, or it may point to a predilection for analyzing specific features of a piece. While this is understandable, given the cognitive burden such consideration would impose, it suggests the lack of a universal or macro-view of musical phenomena and their reception.

Lastly, Table 4.4.3 lists 71 of 112 Types with terminal slot 'of'. Bundles not falling into the categories of numeratives or disciplinary content (e.g., ExCo 12 'in terms of) have been excluded. Various Types of numerative bundles account for 61% of all terminal 'of' LB. While there are bundles containing disciplinary terms among these, such terms are contained within numerative bundles. This is the case with nine of the Types; these contain the disciplinary terms: member(s), movement, music, performance, sound, style, title (highlighted in pink in Table 4.4.3). Chart 4.4.4 shows the percentage of each function among these LB, with the portion Type accounting for the largest percentage (37%), which again points to

the descriptive practice of the Discourse that focuses on portions of

a musical piece.

Table 4.4.3: Terminal Slot 'of' (71 Types)

Rank	Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Type
30	a	series	of			aggregate
148	a	set	of			aggregate
243	a	group	of			aggregate
40	the	music	of			facet
76	the	course	of			facet
101	the	develop-ment	of			facet
219	the	sound	of			facet
248	the	effect	of			facet
256	over	the	course	of		facet
511	the	name	of			facet
524	in	front	of			facet
534	the	character	of			facet
536	the	style	of			facet
537	the	title	of			facet
26	a	sense	of			make-up
38	the	idea	of			make-up
49	the	notion	of			make-up
85	the	sense	of			make-up
86	the	concept	of			make-up
102	the	form	of			make-up
125	the	nature	of			make-up
178	the	basis	of			make-up

Table 4.4.3: Terminal Slot 'of' (71 Types)

Rank	Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Type
203	a	form	of			make-up
287	in	the	form	of		make-up
410	the	meaning	of			make-up
412	the	spirit	of			make-up
441	the	act	of			make-up
442	the	face	of			make-up
474	the	issue	of			make-up
475	the	value	of			make-up
3	the	end	of			portion
15	at	the	end	of		portion
16	the	beginning	of			portion
22	a	number	of			portion
46	as	part	of			portion
74	at	the	beginning	of		portion
114	the	start	of			portion
115	the	rest	of			portion
126	the	opening	of			portion
145	second	half	of			portion
152	the	second	half	of		portion
202	the	number	of			portion
258	first	half	of			portion
280	the	majority	of			portion
290	the	first	half	of		portion
307	the	first	of			portion
320	at	the	start	of		portion
375	the	middle	of			portion
452	in	the	second	half	of	portion

Table 4.4.3: Terminal Slot 'of' (71 Types)

Rank	Slot 1	Slot 2	Slot 3	Slot 4	Slot 5	Type
468	is	part	of			portion
551	the	heart	of			portion
552	the	part	of			portion
110	is	one	of			quantum
183	the	perform-ance	of			quantum
226	a	member	of			quantum
266	as	one	of			quantum
285	first	movement	of			quantum
291	the	first	move-ment	of		quantum
339	an	example	of			quantum
376	was	one	of			quantum
399	the	members	of			quantum
517	a	piece	of			quantum
41	a	kind	of			variety
43	the	case	of			variety
66	in	the	case	of		variety
94	a	variety	of			variety
272	a	sort	of			variety
336	the	kind	of			variety
387	this	kind	of			variety
445	a	type	of			variety
460	this	type	of			variety

Having now considered all LB containing the preposition 'of', it is abundantly clear that various types of numeratives (particularly

portions) are such a salient feature of Music Discourse that they constitute a structural feature of it. As such, it is likely that they are also found in other levels of the Discourse.

Chart 4.4.4: Terminal Slot 'of' Functions
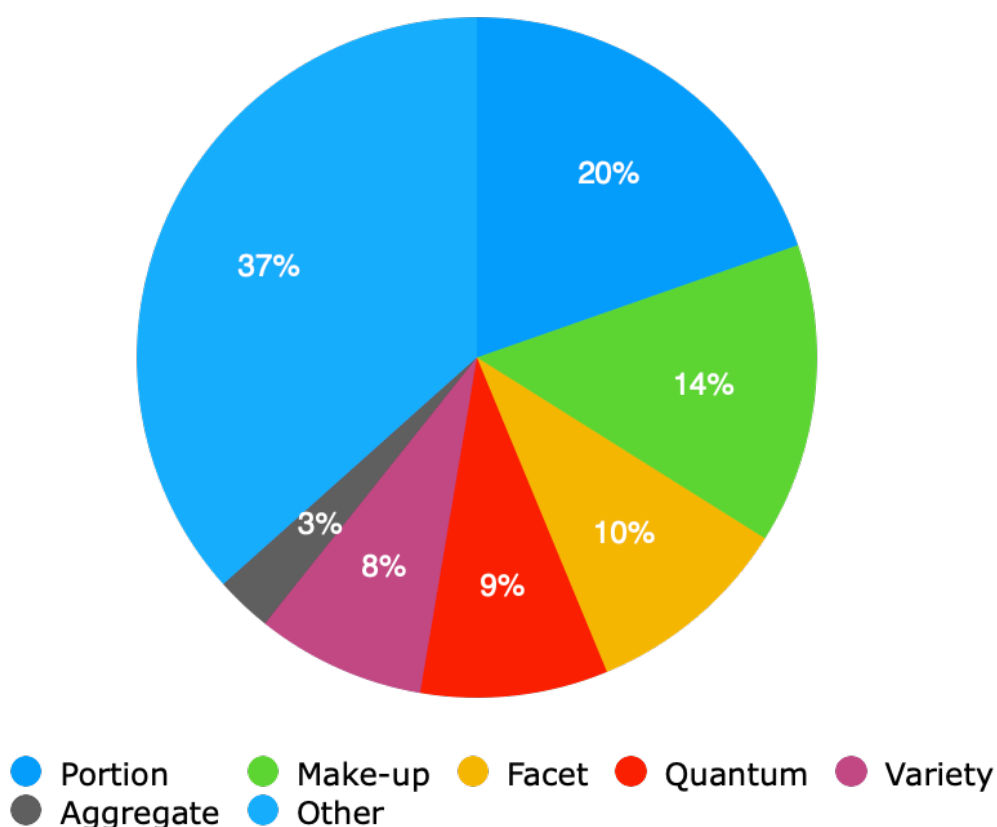
4.5 Extended Lexical Bundles

Given their frequency in the Discourse, numeratives are found in both shorter and longer lexical bundles. The standard range of investigation for lexical bundles has become 3- to 5-LB (Biber et al. 1999; Chen & Baker 2010). Considering that many 3-LB are subsumed under longer bundles, it is likely that a large number of

2-LB would exhibit the same relationship to longer bundles, being subsumed by them. Consequently, study of 2-LB would likely present a considerably less fruitful yet immensely more time-consuming analytical challenge. By contrast, bundles of more than five words might offer insight into specific discoursal features if mined at lower minimums. To test this hypothesis, I searched for LB of six or more words with a minimum frequency of six and minimum range of four. This resulted in 45 Types of 413 total Tokens, many of which are constructed from two or more smaller LB, and all of which were 6- to 8-LB. After eliminating returns that count the possessive 's' as a word and those subsumed under longer extended bundles, a total of 39 Extended LB remain. Of these, eight represent longer discoursal bundles (e.g., 'this is not to say that the') that are not disciplinary-specific; however, the other 31 LB are, as shown in Table 4.5.1 by frequency, range, number of words, number of journals, and function.

Inspection of the Extended LB shows that these 31 are largely focused on portions of pieces (i.e., scores) and historical context; thus, largely numerative in nature. Each bundle's percentage of the total 39 extended LB are shown in Chart 4.5.1. Of course, since these bundles are sometimes drawn from as few as two or three journals, they may be less reliable as indicators of the contents of

Table 4.5.1: 6-, 7-, 8-word LB

Rank	Frequency	Range	Bundles	No. Words	No. Journals	Function
1	36	25	in the second half of the	6	7	portion
2	26	18	in the first half of the	6	7	portion
3	15	8	for the first time in the	6	6	quantum
4	12	10	the turn of the twentieth century	6	6	portion
7	11	6	in the first movement of the	6	4	portion
8	10	5	at the end of the the song	6	2	portion
10	10	9	in the first part of the	6	6	portion
12	10	8	the end of the nineteenth century	6	4	portion
13	10	7	the first half of the twentieth	6	5	portion
14	10	7	the first half of the twentieth century	7	5	portion
15	10	4	the first time in the piece	6	3	quantum
16	9	8	at the beginning of this article	6	5	portion
19	8	7	at the turn of the century	6	5	portion
22	8	5	in the first half of the twentieth century	8	4	portion
23	8	8	late nineteenth and early twentieth centuries	6	6	portion
26	8	7	the late nineteenth and early twentieth	6	5	portion
27	7	5	in the early years of the	6	5	portion
29	7	7	of the work as a whole	6	4	genitive
30	7	7	the late nineteenth and early twentieth centuries	7	5	portion
32	6	5	at the beginning of the second	6	3	portion
33	6	5	at the end of the first	6	3	portion

Table 4.5.1: 6-, 7-, 8-word LB

Rank	Frequency	Range	Bundles	No. Words	No. Journals	Function
34	6	6	at the end of the second	6	5	portion
35	6	4	first half of the nineteenth century	6	3	portion
38	6	5	of the first half of the	6	3	portion
39	6	6	only at the end of the	6	2	portion
40	6	4	second half of the nineteenth century	6	4	portion
41	6	4	the first half of the nineteenth	6	3	portion
42	6	4	the first half of the nineteenth century	7	3	portion
43	6	4	the second half of the century	6	4	portion
44	6	4	the second half of the nineteenth	6	4	portion
45	6	4	the second half of the nineteenth century	7	4	portion

the overall Discourse. The fact that these Extended LB are so consistently and overwhelmingly composed of portion bundles, however, suggests support for the view that Music Discourse relies heavily on numeratives for its narration, both those associated with portions of works/scores and those used to indicate historical context.

Chart 4.5.1: Extended LB Functions
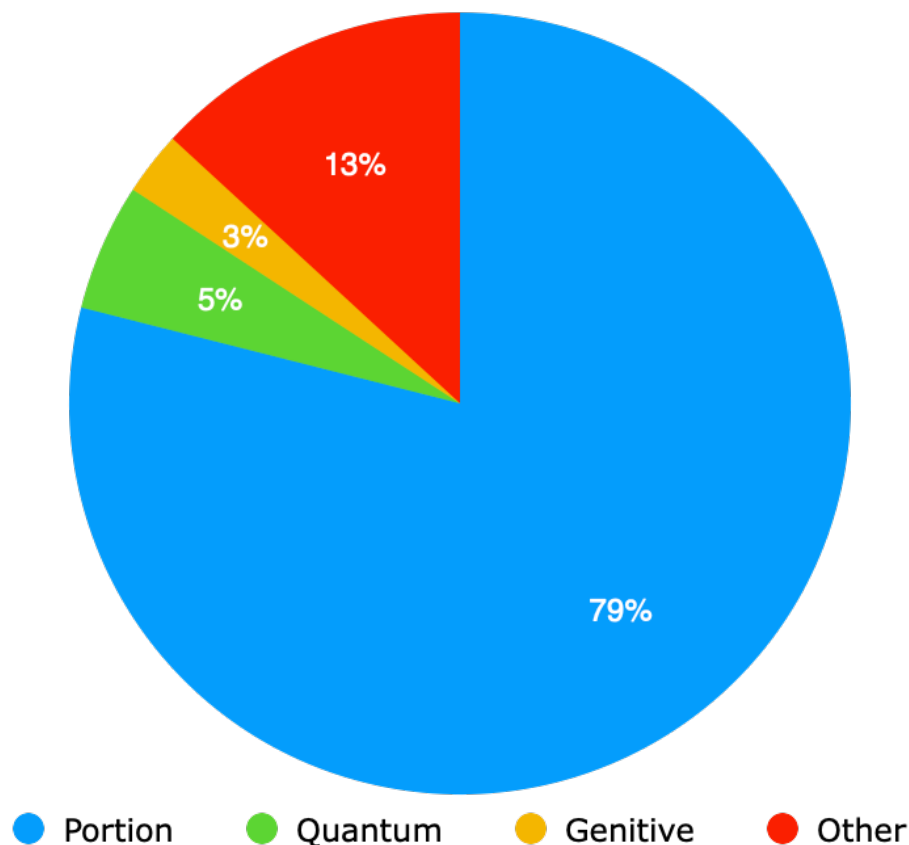
This high percentage of portion bundles draws attention to an interesting feature of the Discourse; namely, a focus on development. For instance, when portions of a piece are discussed, the second half is mentioned nearly twice as often as the first. Chart 4.5.2 maps all of the bundles containing the word 'half'. The chart is numbered and color coded to identify each of these portion bundles. In addition to their ExCo rank (given as a simple numeral), each bundle in the chart is labelled by frequency (F#) and range (R#). Boxes of the same color in the chart belong together as a single LB. In the case of ExCo 44, for example, the word 'half' is the beginning

of the bundle. In four of these instances, there are Types ending in 'of' and also 'of the'. Arrows indicate overlaps between LB and font size is adjusted proportionally by Rank. As can be seen, the majority of these refer to 'second half', again suggesting a teleological view concerned with the purpose of musical development. This in turn may help to explain why the Discourse is concerned with socio-historical context: musical development may be construed as a composer's response to, and microcosm of, the world that gives birth to it. This possibility will be further tested in Section 4.10.

Chart 4.5.2: 'half' as Nodal Word
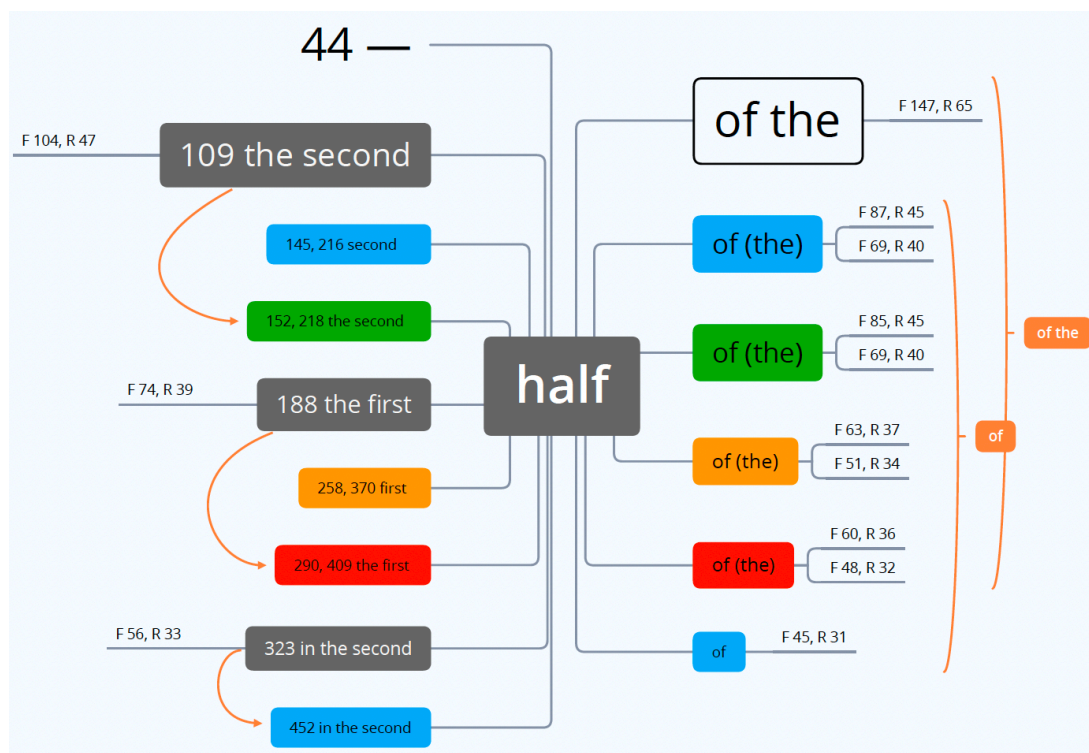
4.6 LB Disciplinary Content

Having established the significance of numeratives to Music Discourse by virtue of their frequency of appearance in lexical bundles, discussion now turns to analysis of the additional nominal content found in these bundles. Excluding the numeratives analyzed above, many of the remaining LB contain nominals that are disciplinary specific, including some that have both a numerative and a disciplinary term (e.g., ExCo 56 'the first movement'). Grouped together with these are several bundles containing the term 'century', which occurs sufficiently often to be considered a part of the propositional content of the Discourse (e.g., 48 'the nineteenth century'). These are included here among the disciplinary LB because they form a distinct set of bundles, a prominence that indicates their importance to the Discourse. After eliminating nominal bundles that are neither disciplinary nor propositional, a total of 69 nominal LB remain, as shown in Table 4.6.1, which also lists their functions. In cases where it may be questionable whether a LB is disciplinary specific (e.g., ExCo 203 'a form of'), the ambient content of each Token has been checked to make that determination. This group of LB support Biber et al.'s (2004) claim that lexical bundles offer a view of semantic content within a given discourse, as well as Qin's (2014) statement that they are disciplinary specific.

Table 4.6.1: LB Disciplinary and Propositional Content

No.	Rank	Slot 1	Slot 2	Slot 3	Slot 4	Function
1	10	of	the	music		genitive
2	24	of	the	piece		genitive
3	29	of	the	song		genitive
4	32	of	the	work		genitive
7	64	of	the	musical		genitive
13	134	of	the	com-poser		genitive
16	170	of	the	theme		genitive
23	214	of	music	and		genitive
27	245	of	the	twen-tieth		genitive
28	245	of	the	nine-teenth		genitive
32	268	of	the	nine-teenth	century	genitive
35	275	of	music	in		genitive
42	306	of	the	nine-teenth		genitive
43	312	of	the	score		genitive
44	344	of	the	nine-teenth	century	genitive
47	351	of	the	century		genitive
56	426	of	the	passage		genitive
60	455	of	the	melody		genitive
63	504	of	the	genre		genitive
10	100	music	in	the		nominal
11	102	the	form	of		nominal
12	105	music	of	the		nominal
14	158	the	right	hand		nominal

211 of 460

Table 4.6.1: LB Disciplinary and Propositional Content

No.	Rank	Slot 1	Slot 2	Slot 3	Slot 4	Function
15	163	the	left	hand		nominal
19	182	structure	of	the		nominal
20	183	the	perform-ance	of		nominal
22	203	a	form	of		nominal
24	219	the	sound	of		nominal
25	221	move-ment	of	the		nominal
26	236	the	music	and		nominal
30	262	music	and	the		nominal
31	264	analysis	of	the		nominal
33	270	section	of	the		nominal
36	281	the	music	is		nominal
45	348	the	music	in		nominal
50	381	interpre-tation	of	the		nominal
52	390	voice	of	the		nominal
55	396	perform-ance	of	the		nominal
58	439	sound	of	the		nominal
62	486	music	as	a		nominal
64	517	a	piece	of		nominal
66	529	music	from	the		nominal
67	536	the	style	of		nominal
68	537	the	title	of		nominal
69	553	the	piece	is		nominal
6	56	the	first	move-ment		por-tional

Table 4.6.1: LB Disciplinary and Propositional Content

No.	Rank	Slot 1	Slot 2	Slot 3	Slot 4	Function
37	285	first	move-ment	of		por-tional
39	291	the	first	move-ment	of	por-tional
8	72	in	the	bass		quali-fying
17	176	in	the	song		quali-fying
18	181	in	the	music		quali-fying
21	198	in	the	form		quali-fying
34	274	in	the	musical		quali-fying
38	287	in	the	form	of	quali-fying
41	298	to	the	music		quali-fying
46	350	in	the	piano		quali-fying
49	371	in	the	piece		quali-fying
51	389	to	the	tonic		quali-fying
53	394	in	the	score		quali-fying
54	395	in	the	work		quali-fying
57	435	in	the	nine-teenth		quali-fying
61	457	that	the	music		quali-fying

Table 4.6.1: LB Disciplinary and Propositional Content

No.	Rank	Slot 1	Slot 2	Slot 3	Slot 4	Function
65	525	in	the	nine-teenth	century	qualifying
5	48	the	nine-teenth	century		temporal
9	93	the	twentieth	century		temporal
29	252	early	twentieth	century		temporal
48	364	the	eigh-teenth	century		temporal
59	454	late	nine-teenth	century		temporal
40	295	be	heard	as		verbal

The above 69 bundles are assigned to one of six categories based on their function within the clause or clause complex as determined by individual word functions, particularly those of the initial and terminal slots, thus: 1) Genitive LB with initial 'of' (e.g., ExCo 10 'of the music'); 2) Nominal LB containing a disciplinary term (e.g., ExCo 100 'music in the'); 3) Portional LB referencing part of a musical piece (e.g., ExCo 56 'the first movement'); 4) Qualifying LB initiating a prepositional phrase (e.g., ExCo 72 'in the bass') or hypotactic clause (e.g., ExCo 457 'that the music'); 5) Temporal LB referencing a century (e.g., ExCo 48 'the nineteenth century'); 6) Verbal LB consisting exclusively of a verbal group, the sole instance being ExCo 295 'be heard as'. These categories have been assigned

to highlight the various functions served by disciplinary-specific bundles. In the case of genitive bundles, determination of function is based solely on the initial 'of', whereas nominal bundles were determined as those containing a disciplinary-specific term that included neither a portion or century, nor that serves as a qualifying element. Thus, the categories of Genitive, Portional, Qualifying and Temporal are sub-categories of Nominal. The sole remaining category, Verbal, includes the only example of a disciplinary-specific bundle with propositional content construed as a verbal rather than a nominal.

Many of the disciplinary and propositional LB repeat terms from other such bundles, particularly those containing the word 'music'. An overview of these disciplinary terms by the proportion in which they appear out of the total 69 bundles is shown in Word Cloud 4.6.1. (N.B.: Compound nominals had to be hyphenated in order to appear correctly and at the correct proportion in the cloud.) Among these 69 LB, there are five 4-LB, three of which refer to centuries, while another refers to form and the remaining is a qualifying expression of musical form. The order in which the 64 3-LB appear by frequency provides an overview of the Discourse's epistemological concerns. Thus, bundles 1-4 move from music as phenomenon to specific manifestations of it: piece, song, work.

Word Cloud 4.6.1: Disciplinary and Propositional Content Terms

performance
structure
bass **piece** style twentieth
section
hand eighteenth-century century
nineteenth musical passage voice
sound composer piano
heard **music** analysis
score genre tonic title form
nineteenth-century interpretation
work song theme
melody movement
twentieth-century

Bundles 5-13 introduce temporal context (ExCo 48 'the nineteenth century' and 93 'the twentieth century'), form (ExCo 56 'the first movement' and 11 'the form of'), harmony (ExCo 72 'in the bass'), and creative agent (ExCo 134 'of the composer). These first thirteen disciplinary bundles offer a particularly intriguing view of the Discourse because they outline a discoursal procedure: A) choose a piece/song/work to discuss; then B) contextualize it historically

through an analysis of its harmonic structure. This aligns with the idea that discourse communities have agreed standards of practice (Peregrin 2012:210). If the above procedure seems an oversimplification, consider the sequence of the remaining bundles as shown in Diagram 4.6.1. The presence of numerous nominal bundles, both disciplinary specific and numerative is an indicator of the strength of the present study's methodology for identifying disciplinary content.

In the Diagram, LB disciplinary terms are presented in the order in which they appear by frequency, from the top down, and from left to right. Where terms can be grouped, they are presented as subtopics of a more general nominal that appears before them in this order. Where generic categorical nominals are missing, they have been supplied in brackets (e.g., compositions, historical context, form, harmony). The first and last instances of 'analysis' are also bracketed because the actual term appears at a central position: 31 of 69. Though 'title' appears nearly last, it has been attached to 'score' as a sub-sub-branch.

Simply by frequency, then, the order of propositional content is: music, composition, historical context, form, harmony, composer, performance, voice, and genre, with their respective sub-branches.

Diagram 4.6.1: Sequence of Disciplinary LB
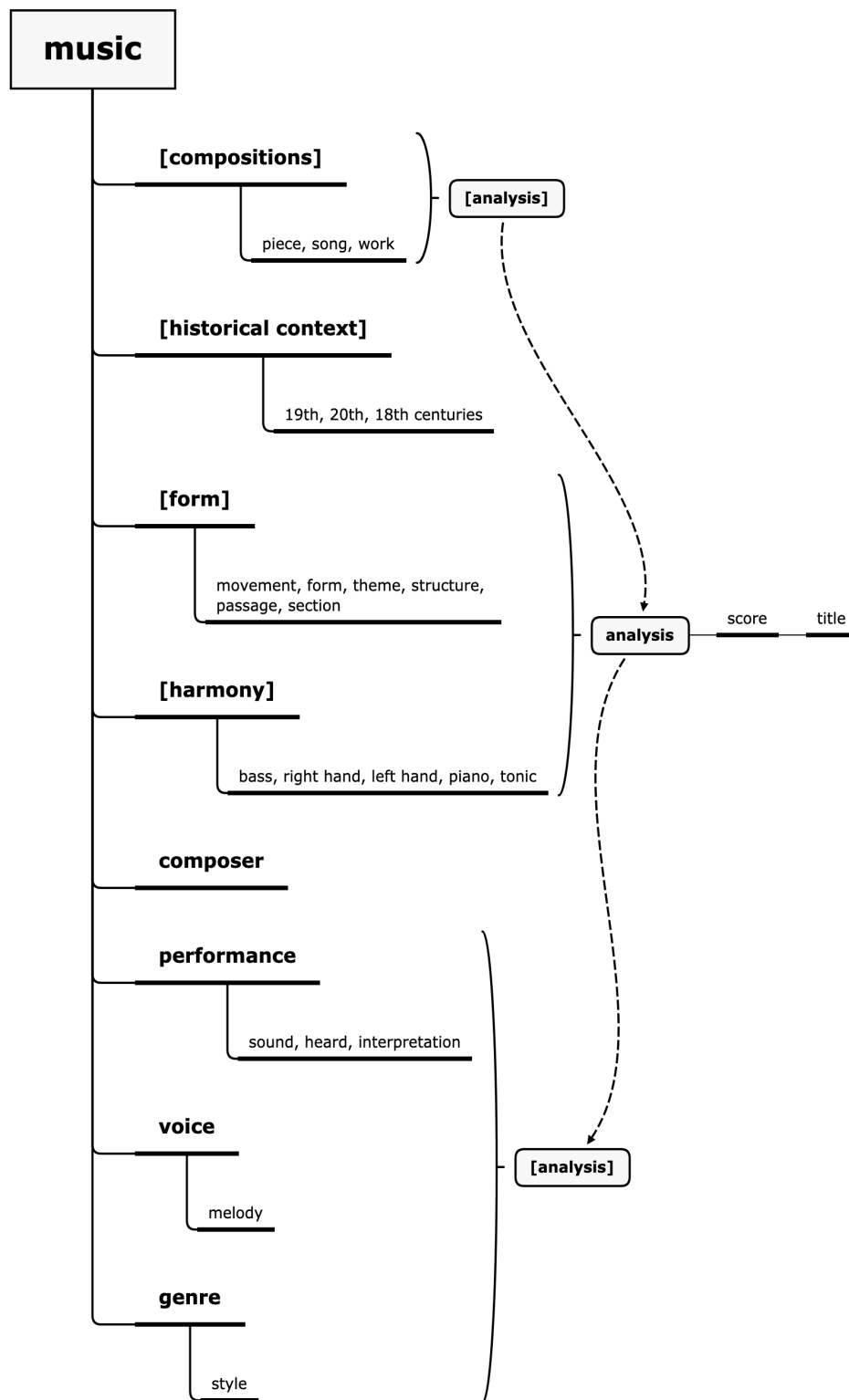
Though compositions are prioritized (by frequency) above all else it is nonetheless remarkable that historical context is the next priority, most notably the Nineteenth, Twentieth, and Eighteenth Centuries, in that order. Clearly, the Discourse is concentrated on a limited historical span. This fact is even more evident from Table 4.6.2, which lists all *century* LB by frequency and range. This view demonstrates an overwhelming preoccupation with the nineteenth and twentieth centuries, suggesting a further concentration on the liminal space spanning the latter part of the nineteenth century and early part of the twentieth. Even by range, the *nineteenth century* bundles appear across a minimum of 39% of ExCo texts. *Twentieth century* bundles are nearly equally represented. Compare this to the single *eighteenth century* bundle, which appears in only 18% of ExCo texts. All of this demonstrates a concentration in the Discourse on music of the so-called Common Practice Period, spanning the seventeenth to early twentieth centuries (Harmony 2024), which in turn indicates a highly Western-centric discourse.

Continuing down Diagram 4.6.1, there is also an evident analytical focus on form and harmony, or more likely, form as construed through harmony. Several portions of a piece are listed under form, including movements and sections within them. Most telling, however, is the relatively high frequency of 'bass' and the two hands

of the piano part, as well as the piano part as a whole. This is particularly intriguing, given that these bundles not only appear in proximity to 'analysis' but also well before 'voice' (which may be used in reference to a harmonic element) and 'melody' by rank. Again, it is clear from these frequencies that harmony is substantially prioritized above melody. Also noteworthy is the complete absence of any mention of rhythm, itself a fundamental element of music.

Table 4.6.2: LB containing 'century'

Rank	Lexical Bundle	Frequency	Range
48	the nineteenth century	145	62
93	the twentieth century	112	59
252	early twentieth century	64	38
268	of the twentieth century	62	43
344	of the nineteenth	54	36
351	of the century	53	28
364	the eighteenth century	52	28
454	late nineteenth century	45	28
525	in the nineteenth	41	22

The inescapable conclusion from all of the above is that Music Discourse is preoccupied with analysis of harmonic form, as evidenced by its focus on the piano, the instrument arguably most associated with nineteenth century harmony and harmonic structure. This in turn reinforces the inference of its dependence on

notation as Common Practice Harmony is most easily represented through staff notation represented in scores; thus, the Discourse remains tethered to music of a specific time and place, which is often denoted as the canon of Western Music. Evidently, Musicology is still very much concerned with musical works, which refutes the claim that Musicology has moved beyond such a canonical view. This claim is exemplified by the following quotation, which appears in one of the Ethnomusicology articles of the Expert Corpus:

> Musicology, ethno- and otherwise, has pushed back hard against the latter of each of these pairs ["perception as opposed to intrinsic value"; process contrasted with object, or constructionist as alternative to materialist]; saying good riddance to the suffocating space of the fixed musical work, the biases of representation, the autonomy of musical sound. (Tenzer 2015:23)

Unfortunately, the view of Music Discourse from frequency, formulaicity and content does not concur with Tenzer's assessment. On the contrary, findings from the present research demonstrate that the journals from which this Corpus is built — highly ranked, oft-cited, broad in scope, prestigious — perpetuate a Discourse which runs in the opposite direction of that claimed by Tenzer. This

then holds implications for the discipline, suggesting that Music Discourse continues to depart from the same set of notions and conceptions it has held from its inception in the nineteenth century. In other words, it struggles to escape the gravity of its origins.

While this conclusion will likely not shock anyone familiar with musicological writings, it will confront them with concerns about diversity and representation. Indeed, the latter concern is also evident among those LB containing personal pronouns and proper noun, and all of the pronouns are shown by the number of bundles that contain them in Graph 4.6.1. These pronouns constitute an extremely small fraction of the total number of LB and are thus presented here only in proportion to one another, not to the Corpus. After the authorial 'I', the dominance of 'he' among LB containing personal pronouns indicates once again a focus on Western Music as a canon created primarily by men. Based on the frequency order in which *century* bundles appear, these were principally men of the nineteenth and twentieth centuries. The sole feminine pronoun is the objective/possessive 'her'. From this perspective, men are frequently discussed as agents in the Discourse, though not women. This is inevitable for a Discourse focused on the canon of Western Music, given that it was largely created by European and (later)

North American male composers. Nonetheless, it is striking that this is still the case in the present Discourse.

Graph 4.6.1: LB containing Personal Pronouns
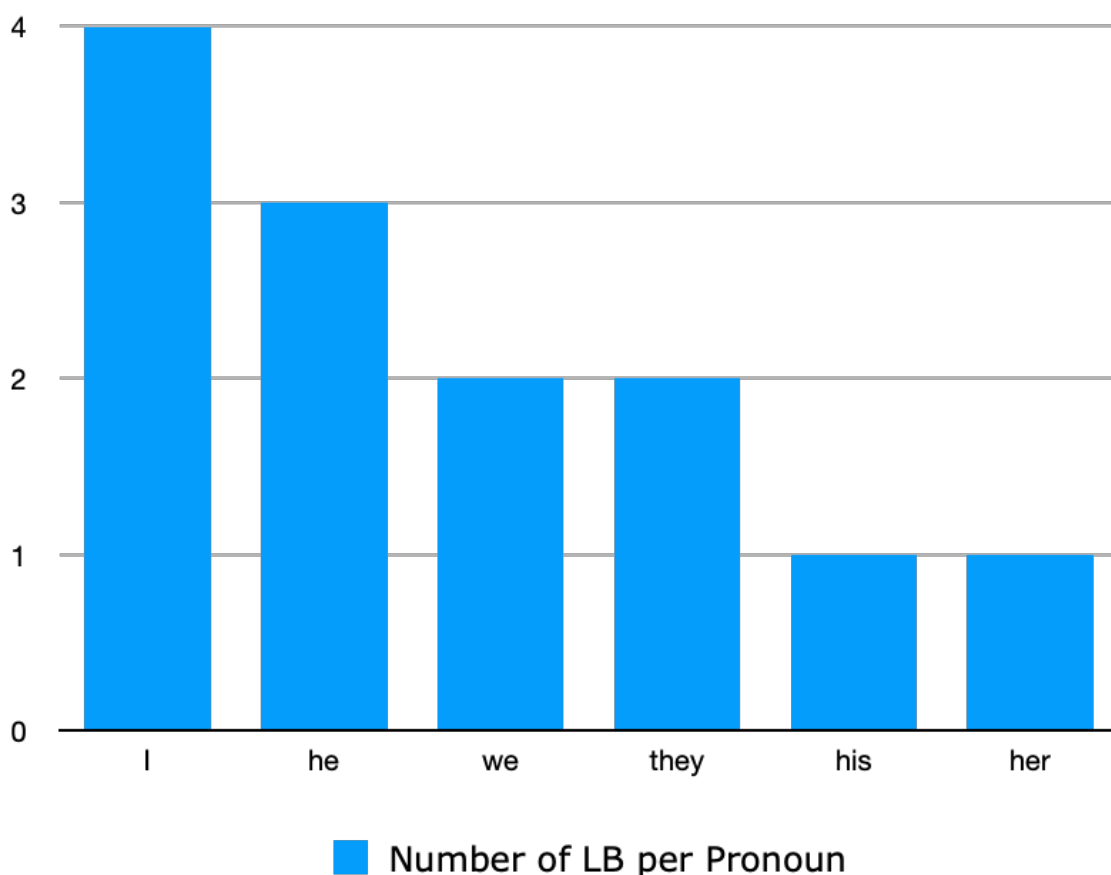
Following personal pronouns, proper nouns also contribute to the Western-centric character of the Discourse. From the general to the specific, these are the United States and New York. Table 4.6.3 lists the five bundles containing complete or partial geographic locations, countries and city. ExCo 227 'in the United' most often is concluded with 'States' yet four Tokens are followed by 'Kingdom'. While this may lend some credence to the argument that journals published in

the United States and the United Kingdom tend to be self-referential, the fact that only four bundles reference the United Kingdom renders this a weak argument. The remaining four geographic bundles are all focused on the United States (ExCo 58) and New York (ExCo 174, 220), specifically, again suggesting that the West is seen at least as a center of musical activity. The large frequency and range of ExCo 58 'the United States' also demonstrates a rather Western-centric view, as it positions America centrally in the Discourse.

Table 4.6.3: LB containing Place Names

Rank	Place Bundles	Frequency	Range
58	the United States	136	41
260	in the United States	63	22
174	the New York	79	20
220	in New York	68	27
227	in the United	67	24

4.7 Music Discourse and Score Notation

As demonstrated above, Music Discourse consists of high-frequency LB among which are multiple numerative bundles and those containing disciplinary terms. That content, organized by frequency, reveals a focus on music of the Western canon, which was largely composed by European and North American men from the

eighteenth through the early twentieth centuries. At this stage of analysis, it should be increasingly clear that the reason for this is that Music Discourse is dependent upon, and thereby delimited by, another semiotic: the musical score. Modern staff notation is a useful tool for representing compositions of significant duration or harmonic complexity and is thus widely accepted as a kind of lingua franca of its own. Since this notation is specific to particular histories of music belonging to a few cultures and arguably even fewer socio-economic classes, any discourse dependent on it is unavoidably confined to those histories, giving little opportunity to engage with other musics. This will be evident in the following section where global views of complete texts offer a telling comparison between subcorpora within the ExCo. However, this is also evident from a reorganization of the LB disciplinary content into musical groupings. These groupings are shown in Map 4.7.1 and its subsequent magnifications. Disciplinary terms from the disciplinary-specific LB are represented in font sizes proportional to the number of bundles containing them. Hence, 'music' figures most prominently as it appears in 16 LB. Because *century* bundles are significant yet separate from disciplinary terminology, they are positioned as a constellation of their own. Also assigned their own constellation are the few bundles directly connected to performance. All remaining disciplinary terms can be gathered and ordered under the agency of

a composer, as all these terms relate directly to score notation and analysis thereof. Magnifications of these three constellations are provided in Map 4.7.1 Magnifications 1-3 for ease of reading.

Though there is little to say about the first two magnifications, as they only contain a few items in readily recognizable relationships, the third magnification shows the more complex web of relationships among LB disciplinary content relating composer to score, score to form, and form to analysis. A overarching focus on harmonic structure is once again apparent, particularly in the grouping of voice, piano (part), and tonic. This last item unequivocally highlights the preoccupation with tonal music, given that the tonic is the primary chord and therefore tonal center of any piece composed within the tradition of Common Practice Harmony.

At this point, it is necessary to mention a limitation to both the Expert and Novice Corpora that directly relates to the question of music notation. ExCo 112 ('shown in example'), listed above, highlights an important aspect of the Discourse that has been excised from both Corpora; namely, score excerpts. While there are relatively few such score examples in Ethnomusicology articles (studies often focusing on music of aural traditions or music recorded in a semiotic other than staff notation), or in novice

Map 4.7.1: LB Disciplinary Content
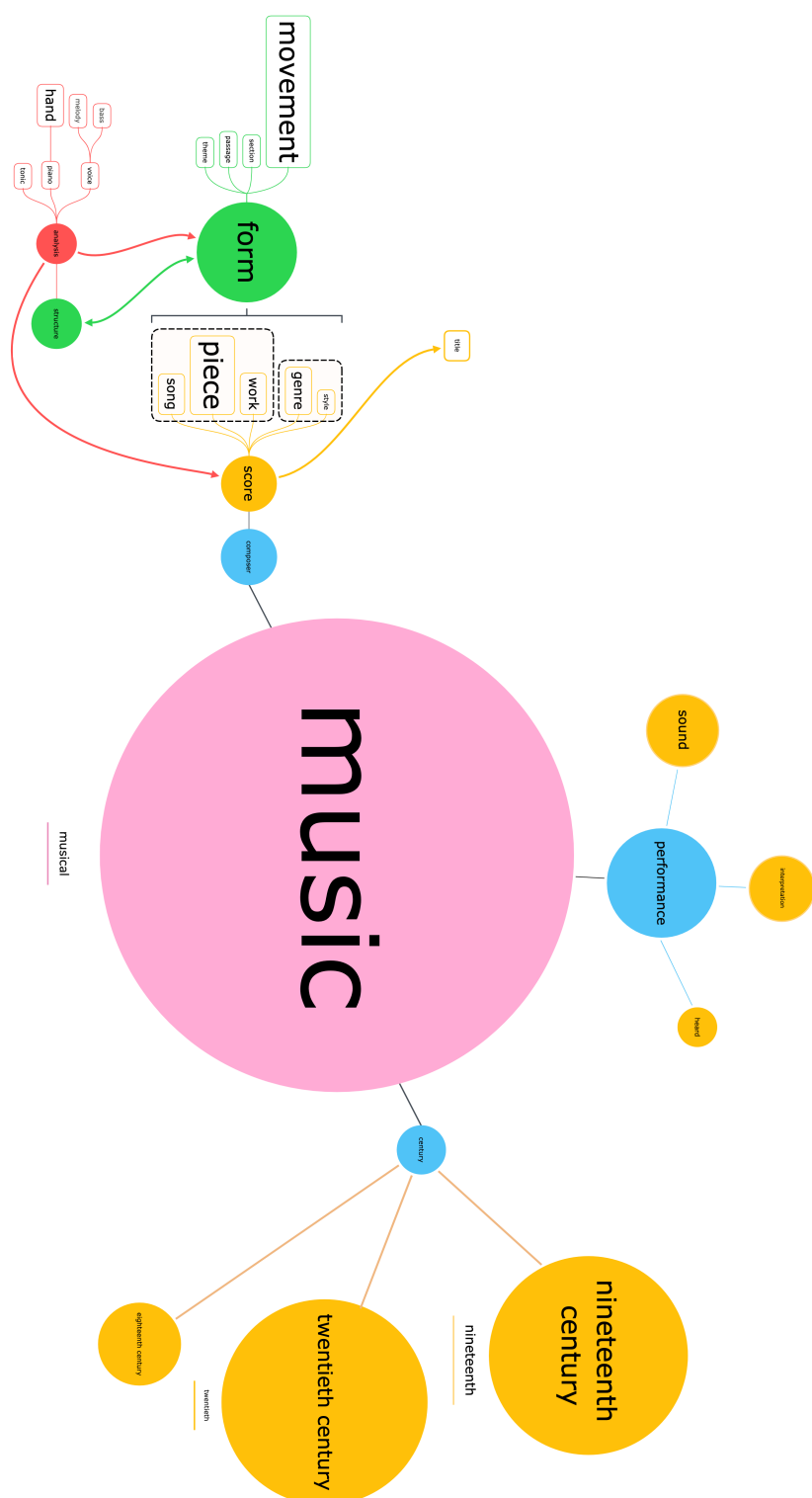
Map 4.7.1 Magnification 1: Century Bundles
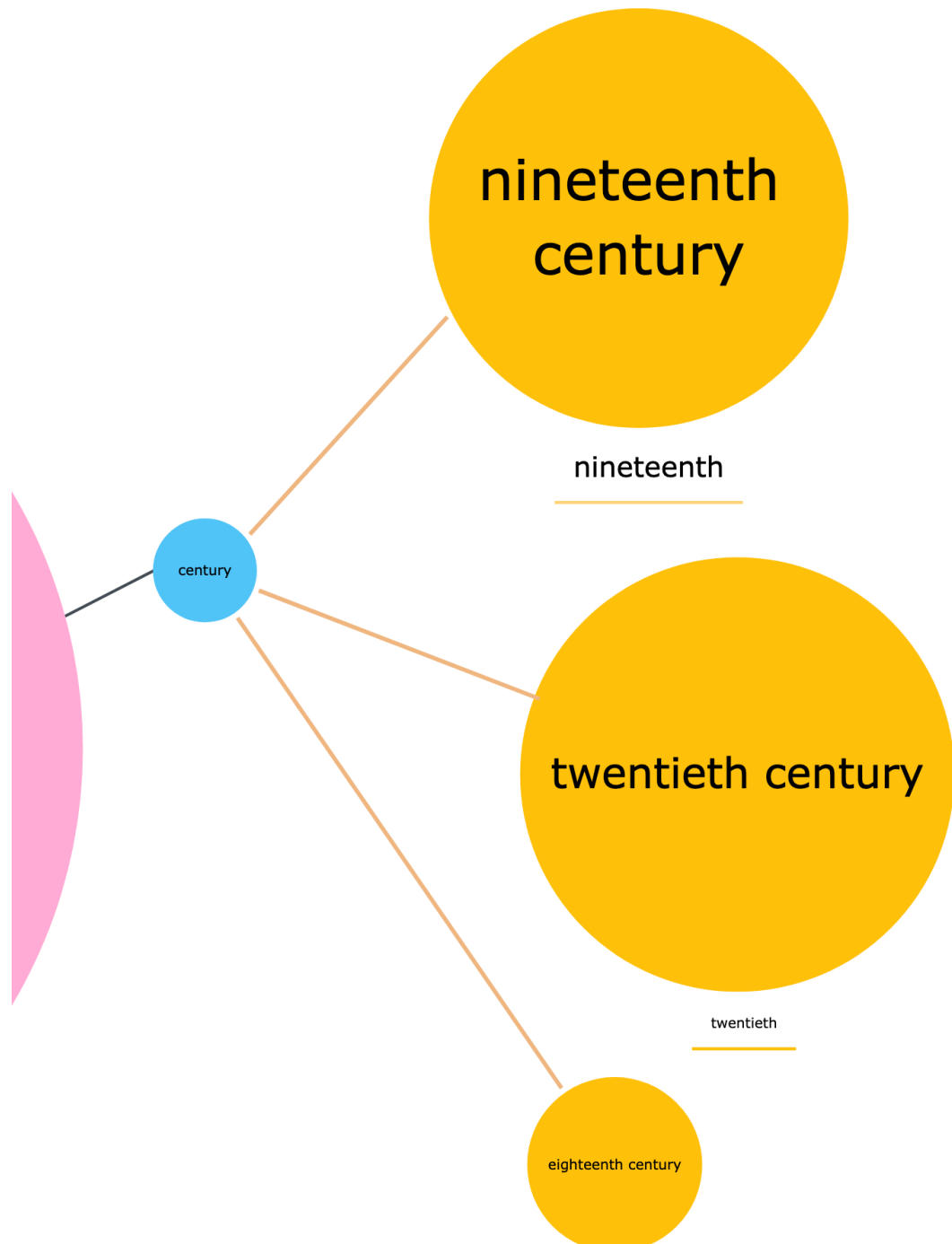
Map 4.7.1 Magnification 2: Performance bundles
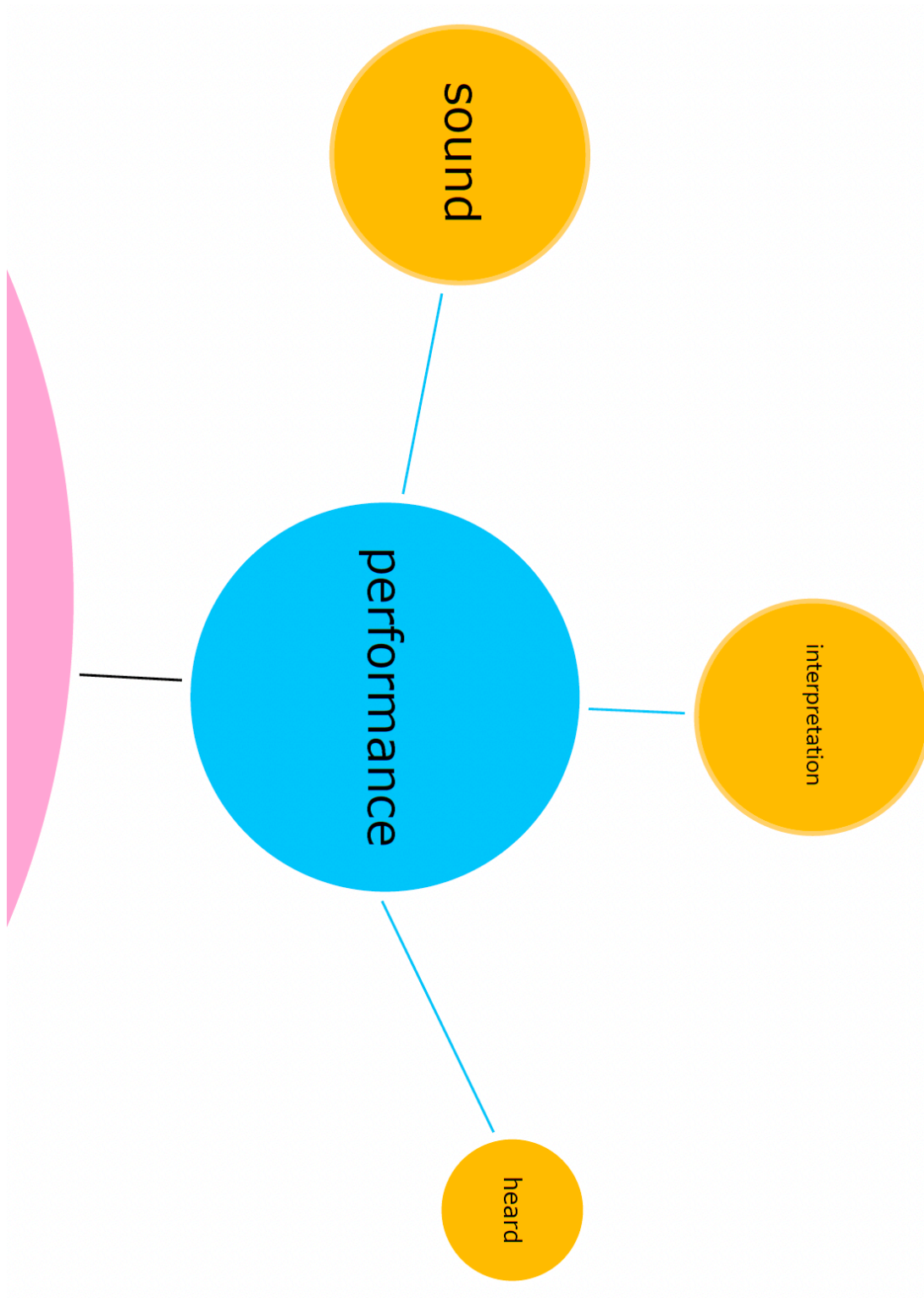
Map 4.7.1 Magnification 3: Form bundles
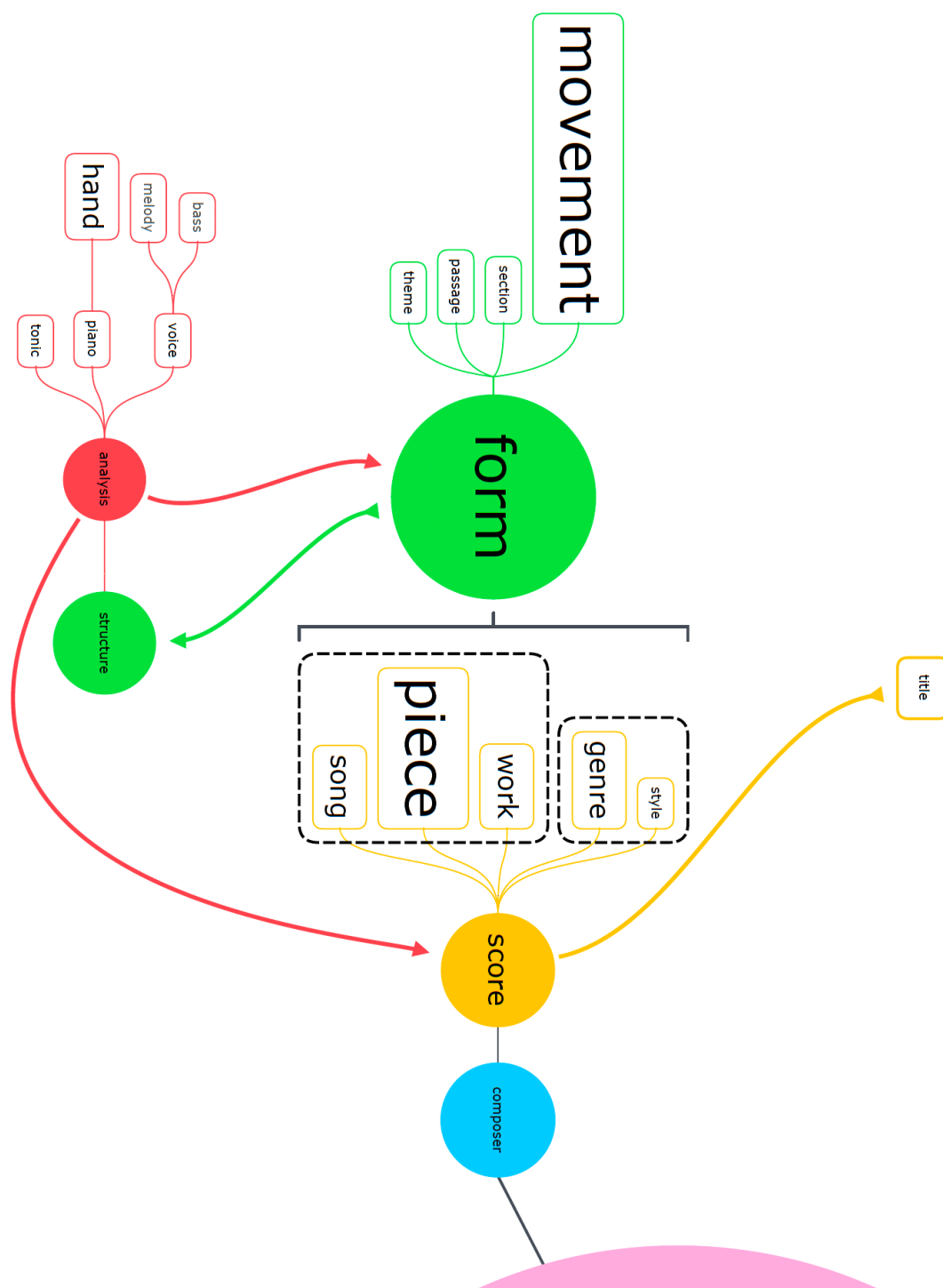
essays, on account of their relative brevity, Musicology articles and especially Music Theory articles rely on such excerpts to illustrate arguments being made in the text. The addition of this primary musical semiotic inarguably forms a fundamental element of the Discourse. Their significance notwithstanding, though, they would not contribute to the methodology of the present research and therefore were set aside for future study.

Table 4.7.1 displays the number of LB containing each of the above disciplinary terms, categorizing them according to their area of Map 4.7.1. These categories are then represented as percentages of the total occurrences of such terms in Chart 4.7.1. Aside from high frequency items, it is noteworthy that several anticipated categories are absent. As mentioned earlier, there is no mention of rhythm. The increasing importance assigned to melody (4%), then harmony (7%), and finally structure (28%) reveals that Music Discourse is predominantly concerned with this last category, affording levels of musical organization the greatest prestige. This view is further enhanced when taking into account the Rank of each of the above bundles (excepting those including the phenomenal items 'music' and 'musical'). Next, Chart 4.7.2 displays these same categories by percentage of the total number of LB Tokens for each term. This view by frequency reveals that time bundles are prioritized. This

Table 4.7.1: Number of LB with Disciplinary Terms

Disciplinary Term	No. of LB	Category	Disciplinary Term	No. of LB	Category
music	16	Phenome-non	piano	1	Harmony
form	4	Structure	tonic	1	Harmony
movement	4	Structure	right hand	1	Melody
piece	4	Structure	melody	1	Melody
nineteenth century	4	Time	voice	1	Melody
twentieth century	3	Time	title	1	Notation
score	2	Notation	heard	1	Perform-ance
Perform-ance	2	Perform-ance	interpretation	1	Perform-ance
sound	2	Perform-ance	passage	1	Structure
musical	2	Phenome-non	section	1	Structure
song	2	Structure	structure	1	Structure
work	2	Structure	theme	1	Structure
nineteenth	2	Time	century	1	Time
analysis	1	Analysis	eighteenth century	1	Time
composer	1	Composer	twentieth	1	Time
bass	1	Harmony	genre	1	Variety
left	1	Harmony	style	1	Variety
left hand	1	Harmony			

focus on chronology suggests a teleological view of music in the Discourse. While it may be argued that structure and chronology warrant this level of attention on account of the former's complexity and the latter's connection to Musicology as a historically oriented discipline, there is no reason why other strata of music should not receive similar attention, from the local phenomena of accent and articulation, to the global concerns of reception.

Chart 4.7.1: Categories of LB Disciplinary Terms

233 of 460
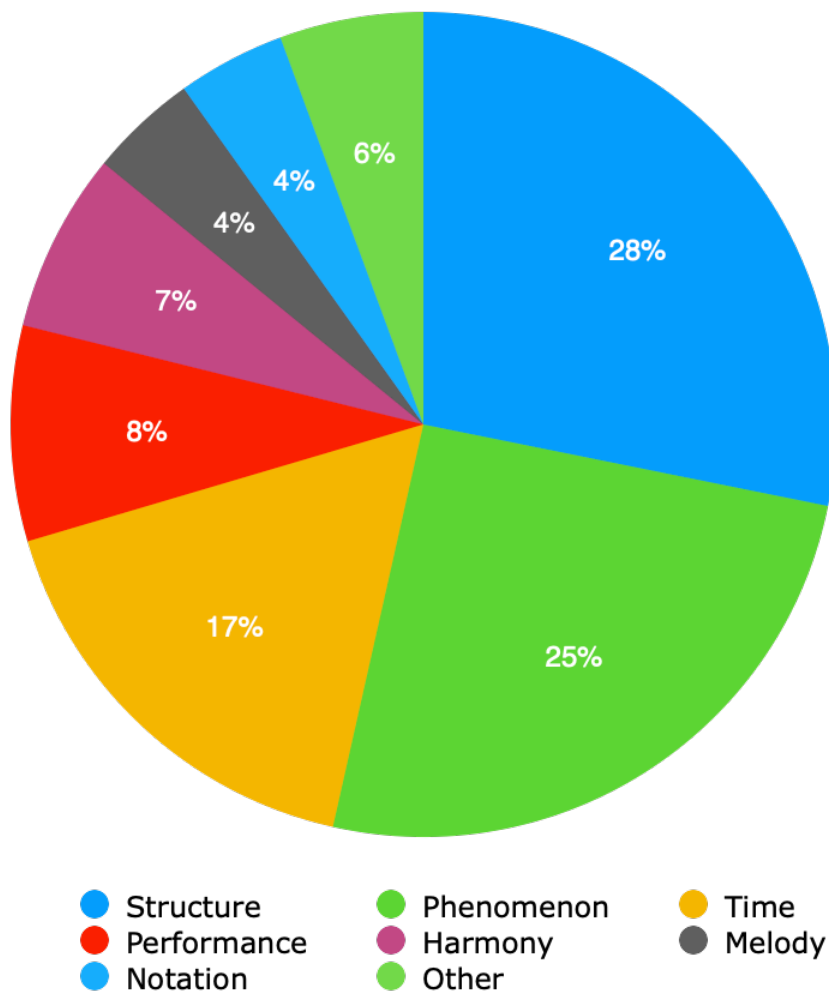
Chart 4.7.2: Categories of LB Disciplinary Terms
(% of Tokens)

Legend: Harmony, Melody, Notation, Performance, Structure, Time, Other
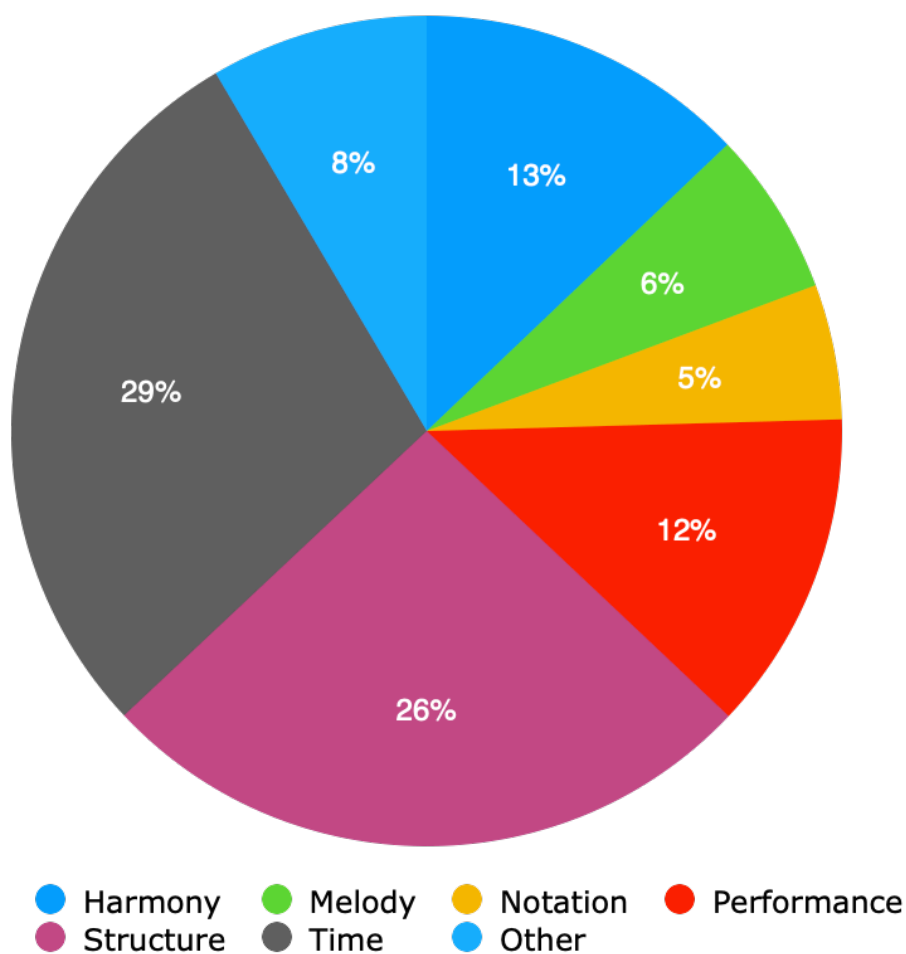
4.8 Global Views of LB in Complete Texts

Next, the disciplinary-specific LB are contrasted with non-disciplinary bundles by location within complete texts. Just as frequency matters, so does place within a text. By reducing font size to three, it is possible to display two complete texts each from all three ExCo subcorpora so as to enable a comparison of lexical bundle usage in Ethnomusicology, Musicology, and Music Theory writings, respectively.

Rather than simply track the frequency and location of all lexical bundles throughout these complete texts, bundles have been color-coded to distinguish those that are non-disciplinary from those that are, the latter category including those containing disciplinary terminology, those that form numeratives, and place and time bundles. To further refine this view of complete texts, quotations have been highlighted in grey, and three large-scale portions of each text have been titled: Introduction, Body, Summary/Conclusion. For journal articles, the determination of these structural boundaries was made by reference to sub-headings within each text. The purpose of this division is to observe whether LB occur with greater or lesser density in any of these three sections. The results are shown in Global Views 4.8.1-4.8.7. To aid the eye, the following color coding scheme is employed in each of the complete texts:

Text

INTRODUCTION

BODY

[QUOTATION]

SUMMARY/CONCLUSION

NON-DISCIPLINARY LB

DISCIPLINARY LB

Global View 4.8.1: Ethnomusicology (Text A)
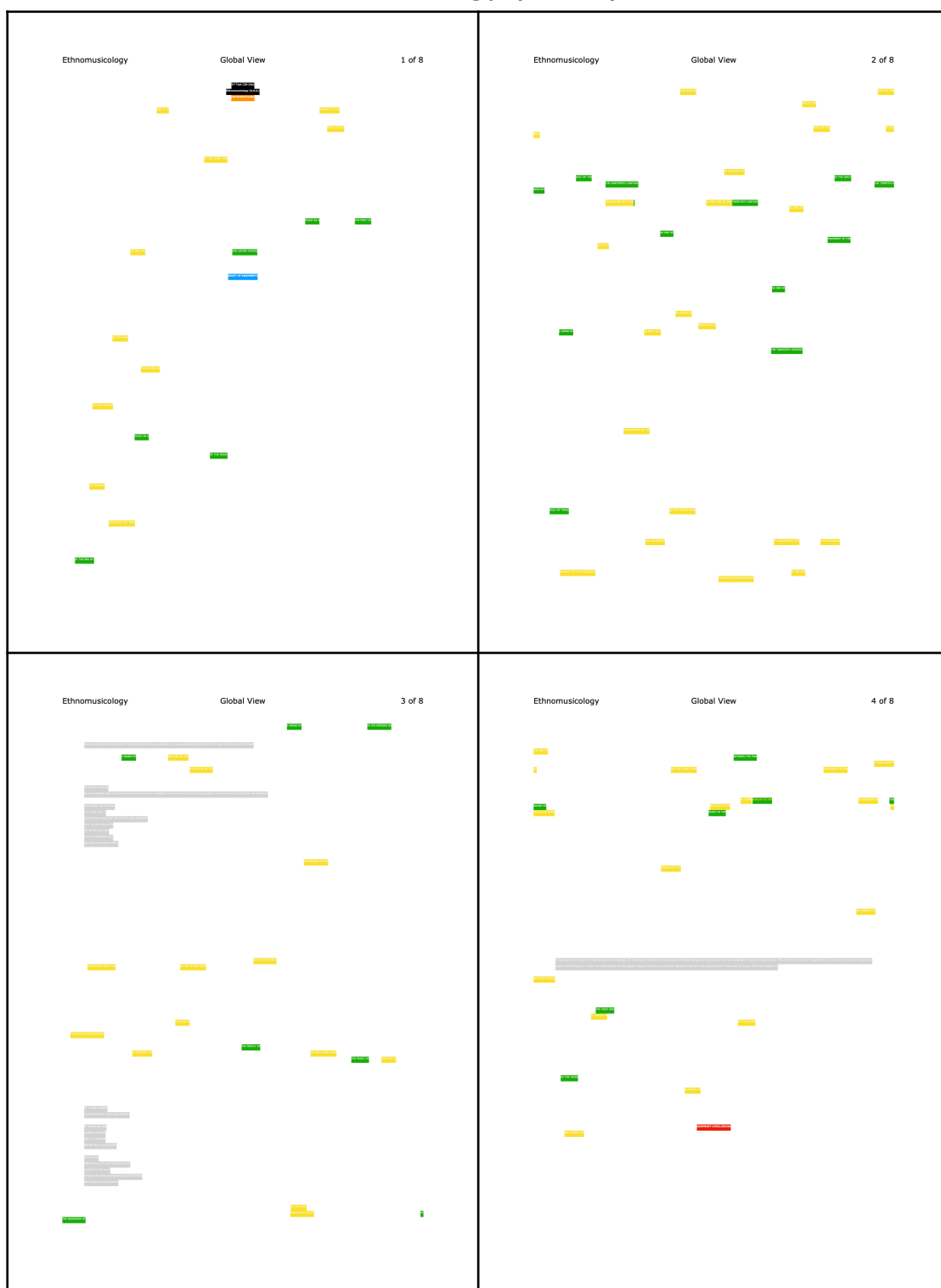
Global View 4.8.2: Ethnomusicology (Text B)
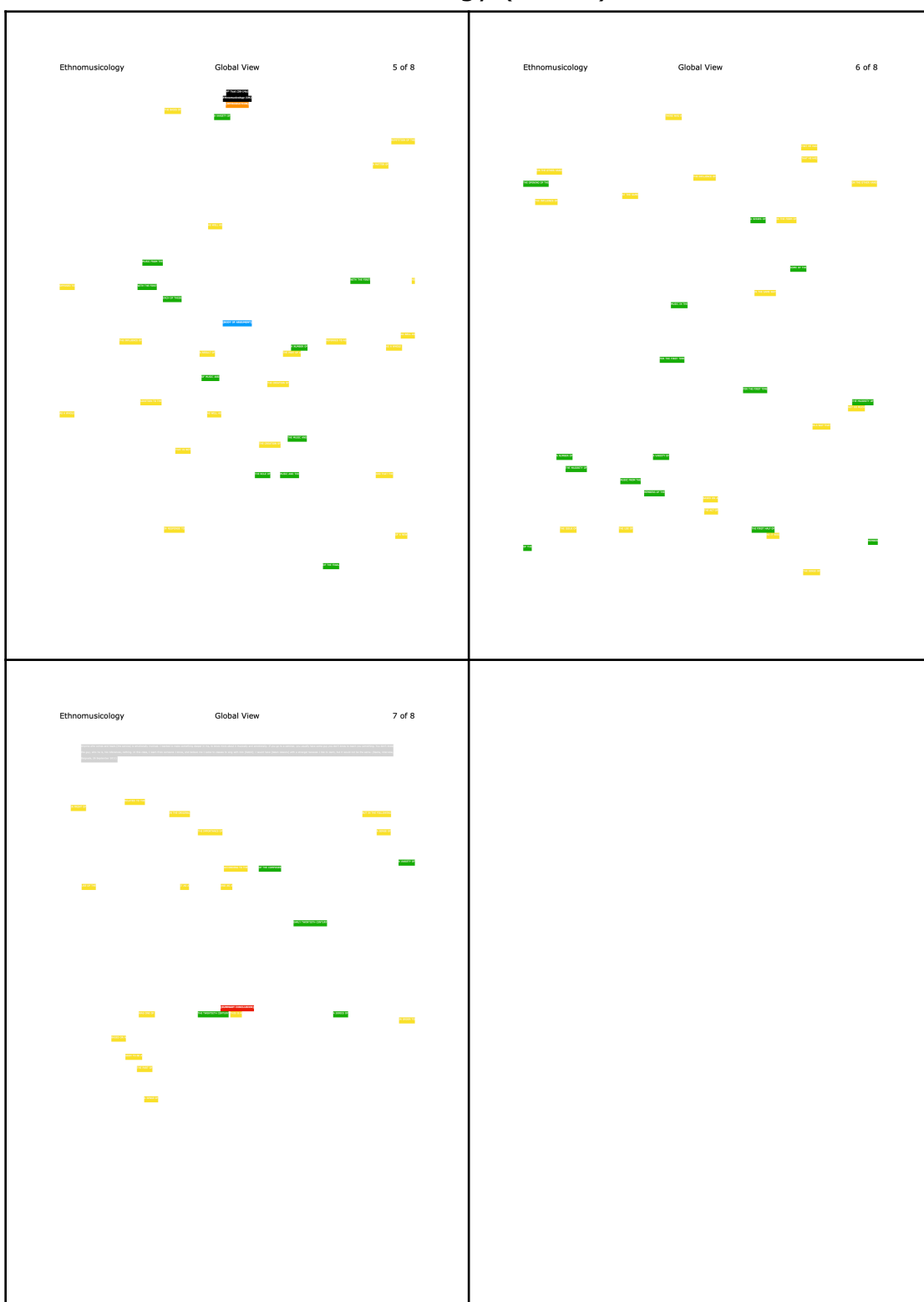
Global View 4.8.3: Musicology (Text A)

Global View 4.8.4: Musicology (Text B)
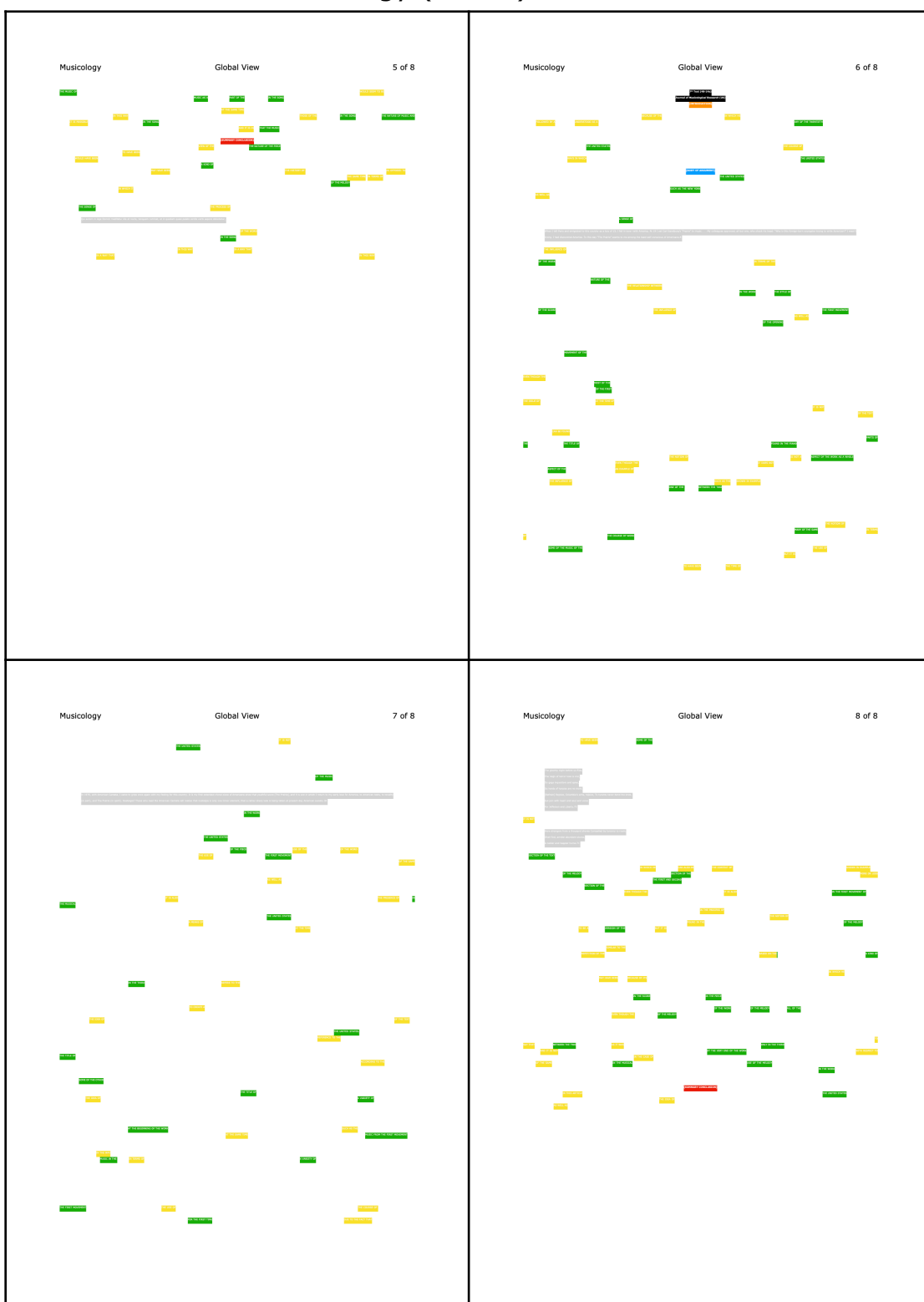
Global View 4.8.5: Music Theory (Text A)
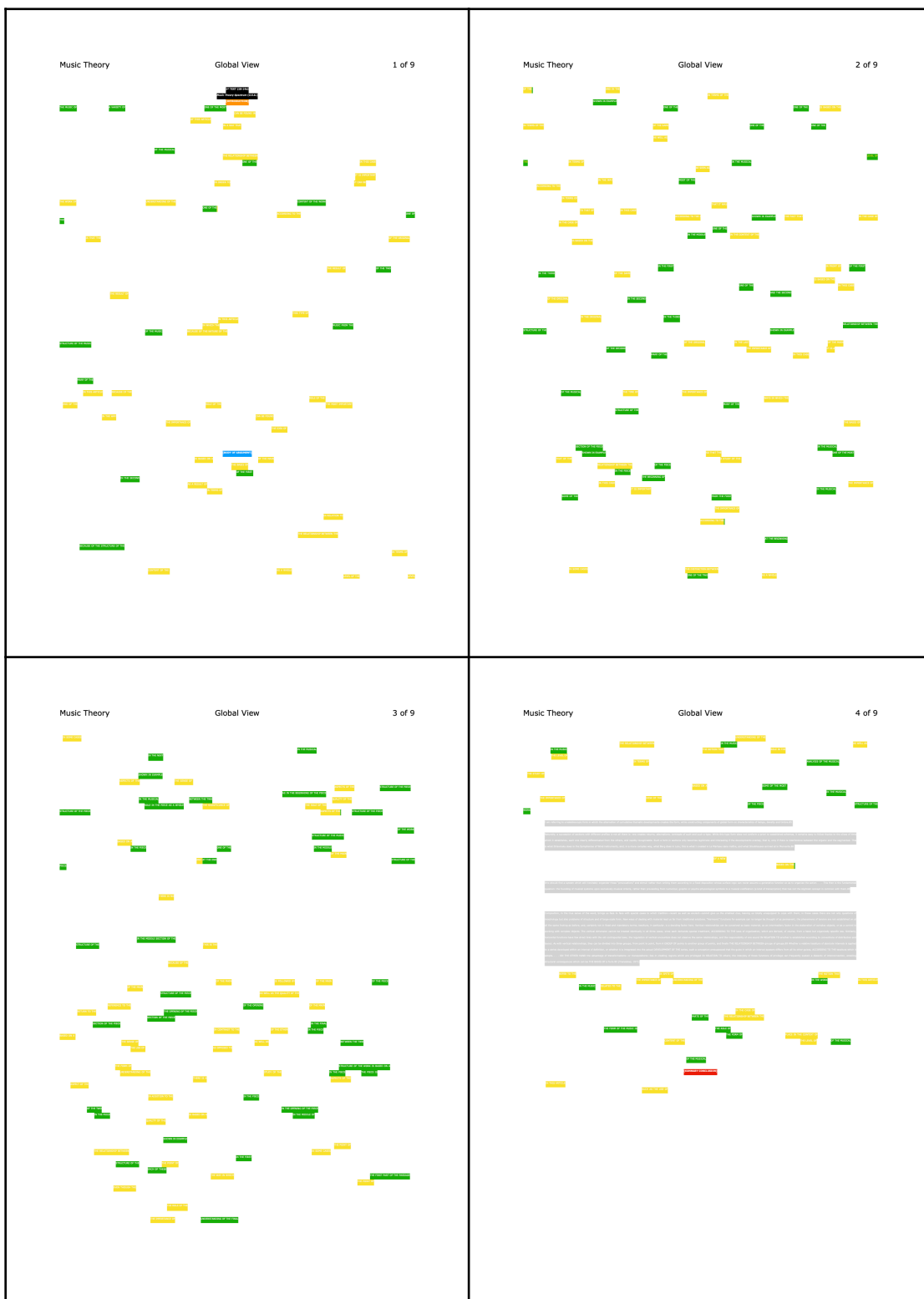
Global View 4.8.6: Music Theory (Text B1)
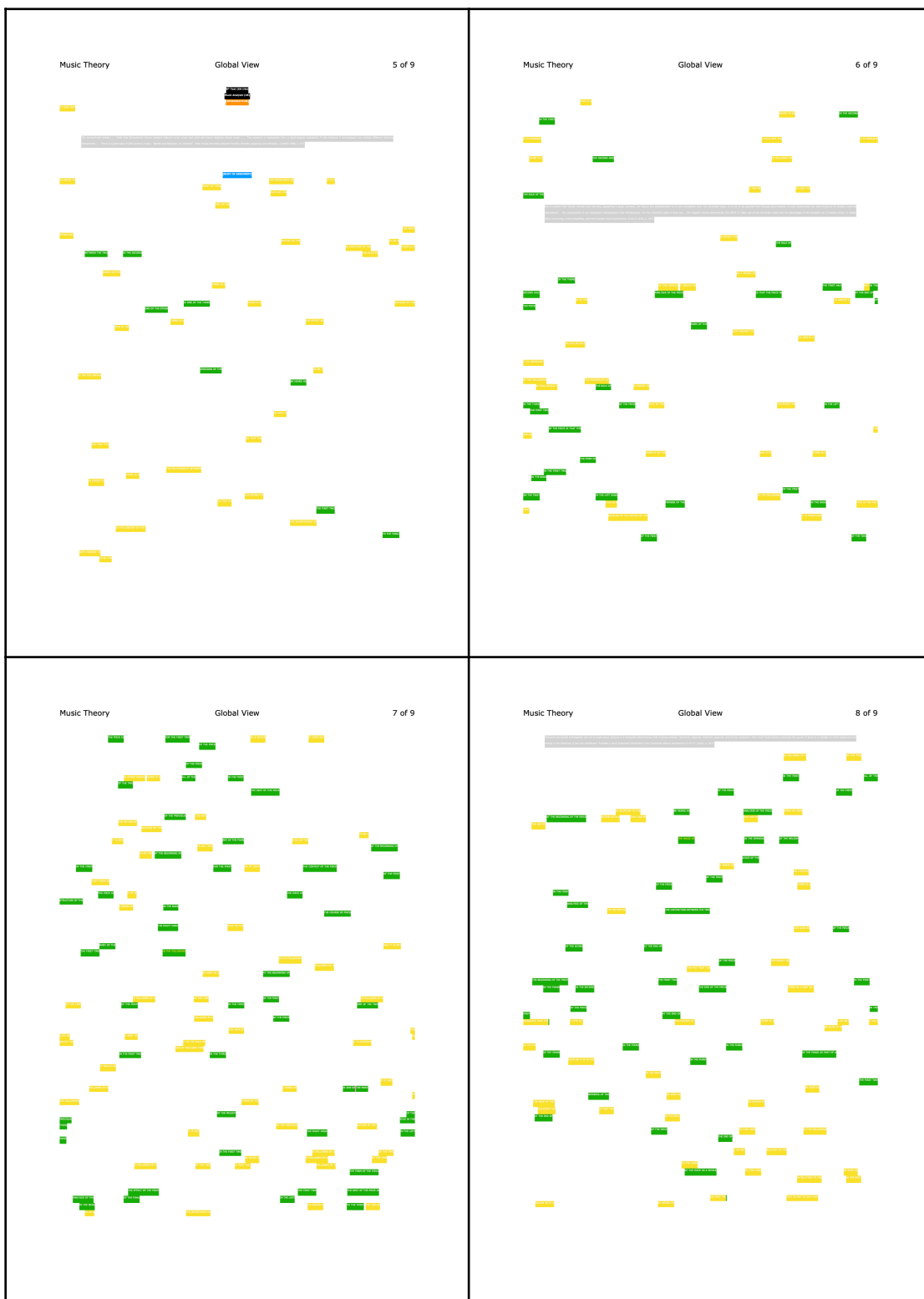
Global View 4.8.7: Music Theory (Text B2)
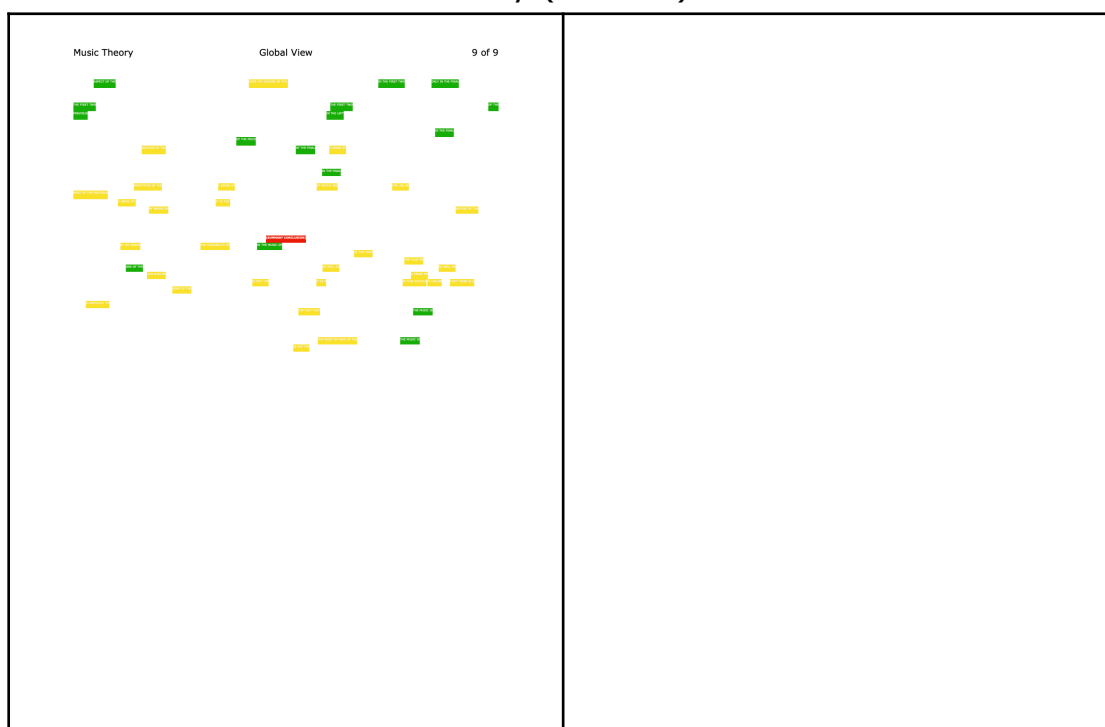
Among the Global Views, the subcorpora are represented in this order: Ethnomusicology, Musicology, Music Theory. Though individual texts cannot represent the whole Corpus, some of the patterns herein are suggestive of each sub-Discourse when compared with the other two.

Beginning with the two Ethnomusicology articles, there is a notable amount of open space — stretches of text in which there is no lexical bundle. Where bundles do appear, non-disciplinary outnumber disciplinary-specific ones. This is unsurprising considering that Ethnomusicology does not rely on Western staff notation nearly as greatly as do Musicology and Music Theory. Thus,

there is no use for the various bundles related to structure, including numeratives, nor much use for the *century* bundles identified above in Table 4.6.2. There are a few quotations, though not many. Overall, use of LB is relatively sparing. In contrast, the two Musicology articles present a rather different vista, with a greater density of bundles, both disciplinary and non-disciplinary, compared to the Ethnomusicology texts. There is also a greater use of quotations. Similarly, the two Music Theory articles display a still greater density of both Types of bundles, though fewer quotations. The only other observation to be made about all six articles is that disciplinary bundles are largely absent from the Summary/ Conclusion sections. This makes sense given that this section would mention works or pieces rather than portions thereof.

4.9 LB Discourse Function

The penultimate stage of analysis involves the application of Biber et al.'s (2004) Lexical Bundle Taxonomy to the ExCo bundles. At the macro level, this taxonomy recognizes three primary discourse functions among LB: discourse, referential, and stance bundles. Because that study did not examine 3-LB, I have added a 'propositional' category to the taxonomy to account for disciplinary-specific bundles. Though each of the three primary categories is further subdivided, those subdivisions are not employed here since

they would not offer greater explanatory power than the four categories listed above. The reason for this is that these macro categories present a view of the Discourse as predominantly dependent on referential bundles: those that introduce or reference propositional content. As the final stage of analysis is a study of the ambient propositional content surrounding each bundle, it is unnecessary to move beyond identification of these macro functions. The purpose of Biber et al.'s (ibid.) study is to categorize the specific functions of bundles within the Discourse rather than to analyze the content of the surrounding slots.

The complete taxonomic analysis is too extensive to list each item here. Instead, Chart 4.9.1 displays the proportions of the four macro categories for the Expert Corpus. By far, referential bundles account for the majority, which is unsurprising for academic writing. This indicates not only that the syntagmatic slots surrounding LB are likely to be rich in propositional content, but moreover that succeeding slots are likely to be more productive of such content than preceding. Examples of these referential bundles include: ExCo 2 'one of the', 117 'of the same', 253 'in relation to the'. These LB are often genitive types that present the Head of a dissociated Thing, as in the first of the preceding examples.

244 of 460

Chart 4.9.1: LB Discourse Function

6% 6% 2%

86%

● Referential ● Discourse ● Stance ● Propositional

BLANK
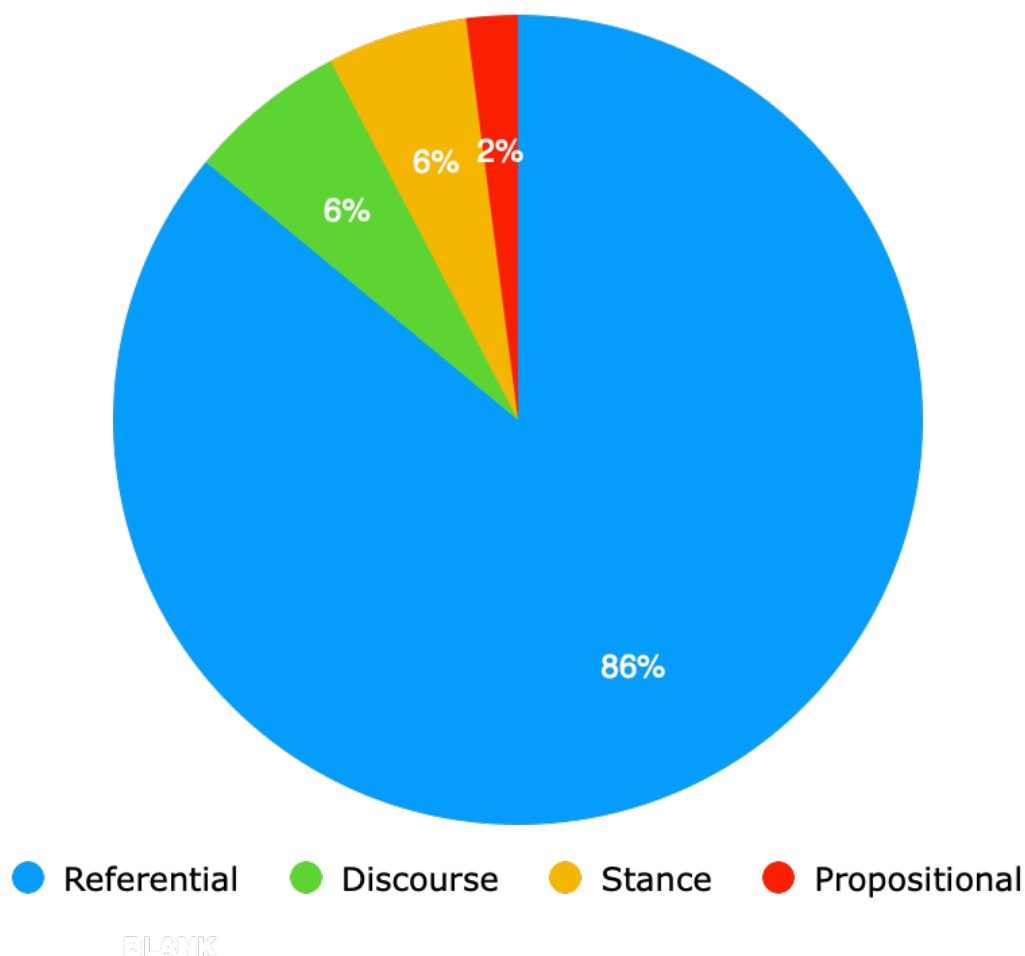
Among the other three categories of bundles, stance is typically represented either by hedging strategies, conditional verbs, or first-person pronouns. Discourse bundles include those LB that relate parts of the Discourse to one another (e.g., ExCo 51 'on the other hand') or that point to something as being within the text (e.g., ExCo 230 'of this article'). Finally, propositional bundles contain disciplinary-specific content, such as the meronym ExCo 163 'the left hand', referring to half of the piano part, or such as 48 'the nineteenth century'.

4.10 Ambient Content

The final stage of analysis maps the ambient propositional content of LB in the Expert Corpus. For this cartographic analysis, all propositional content immediately preceding and succeeding lexical bundles was identified and categorized according to a music taxonomy created for this purpose. While not all preceding (P-slot) or succeeding (S-slot) slots are filled with nominal content, a substantial portion are, as shown in Table 4.10.1. As predicted by the number of referential bundles identified in the preceding stage of analysis, the S-slot is especially productive, with 92% of all slots being filled with propositional content. Several hundred slots are filled with two or more nominal groups paratactically linked by a conjunction. These have been counted separately, resulting in a total of 22,711 items. As this data is far too exhaustive to include here, only the bespoke taxonomy of this study is presented in Table 4.10.2. This taxonomy separates extramusical from musical content. A total of 48 categories account for the latter. Among these, three require clarification: 'composition', '[quotation]', and 'extramusical'. The first includes not only the term 'composition' but also all musical titles referenced in either slot. The second refers to slots in which a single word or nominal group is presented as a discrete item in quotation marks (as opposed to being part of a lengthier quotation). The third refers to all items, generic or otherwise, that do not clearly

246 of 460

fit into any of the musical categories.

Table 4.10.1: Number and Percentage of Productive Ambient Slots

	No. of Filled Slots	% of Total 17,837
Preceding Slot (P-slot)	5,179	29%
Succeeding Slot (S-slot)	16,456	92%

According to the table, 25% (12) of the categories account for 88% of the total 22,711 Tokens. Among these twelve, the top three account for a full 65% of all Tokens. The entire table is visualized below in Chart 4.10.1, where again percentage of the total is represented by proportional font sizes and all categorizes less than 1% of the whole are represented together as '[other]'.

Table 4.10.2: Ambient Content Taxonomy

Category	Frequency	% of Total (22,711)
Extramusical	10,472	46%
Composition	2,586	11%
Structure	1,899	8%
Rhythm	765	3%
Harmony	754	3%
Pitch	738	3%
Analysis	695	3%
Part	633	3%
Melody	533	2%
Performance	455	2%
Musicality	450	2%

Table 4.10.2: Ambient Content Taxonomy

Category	Frequency	% of Total (22,711)
[Quotations]	363	2%
Performer	222	>1%
Notation	219	>1%
Style	200	>1%
Instrumentation	167	>1%
Composer	165	>1%
Instrument	160	>1%
Recording	151	>1%
Counterpoint	149	>1%
Text	130	>1%
Listener	110	>1%
Sound	100	>1%
Musician	78	>1%
Dynamic	68	>1%
Texture	67	>1%
Tempo	50	>1%
Criticism	34	>1%
Ornamentation	34	>1%
Musicologist	32	>1%
Venue	31	>1%
Articulation	30	>1%
Timbre	30	>1%
Emotion	29	>1%
Technique	28	>1%
Range	26	>1%
Critic	22	>1%
Organization	7	>1%

Table 4.10.2: Ambient Content Taxonomy

Category	Frequency	% of Total (22,711)
Trend	6	>1%
Producer	5	>1%
Transmission	5	>1%
Author	2	>1%
Community	2	>1%
Composer/Performer	2	>1%
Editor	2	>1%
Manager	2	>1%
Tuning	2	>1%
Recordist	1	>1%

Though the complete taxonomic analysis is not presented here, a few representative samples from the 11 largest categories can serve to demonstrate how categories were determined (category 12 is excluded here because it consists entirely of quoted material). These samples were extracted by taking the first word to begin with each successive letter of the alphabet, alternating between odd- and even-numbered letters by successive category, thus: 'extramusical' (A, C, E, G, I, etc.), 'composition' (B, D, F, H, J, etc.). These samples are presented in Table 4.10.3. Some categories do not contain certain letters of the alphabet and therefore are represented by fewer samples. To determine whether a word was strictly musical in nature, it was necessary to carefully inspect its context; hence, 'quadrivium' was assigned to the 'extramusical' category on account

Chart 4.10.1: Ambient Content Categories by percentage

extra-
musical

composition

structure

rhythm

harmony

pitch

analysis

part

melody

performance

musicality

[quotations]

[other]

rhythm

harmony

pitch

analysis

part

melody

performance

musicality

[quotations]
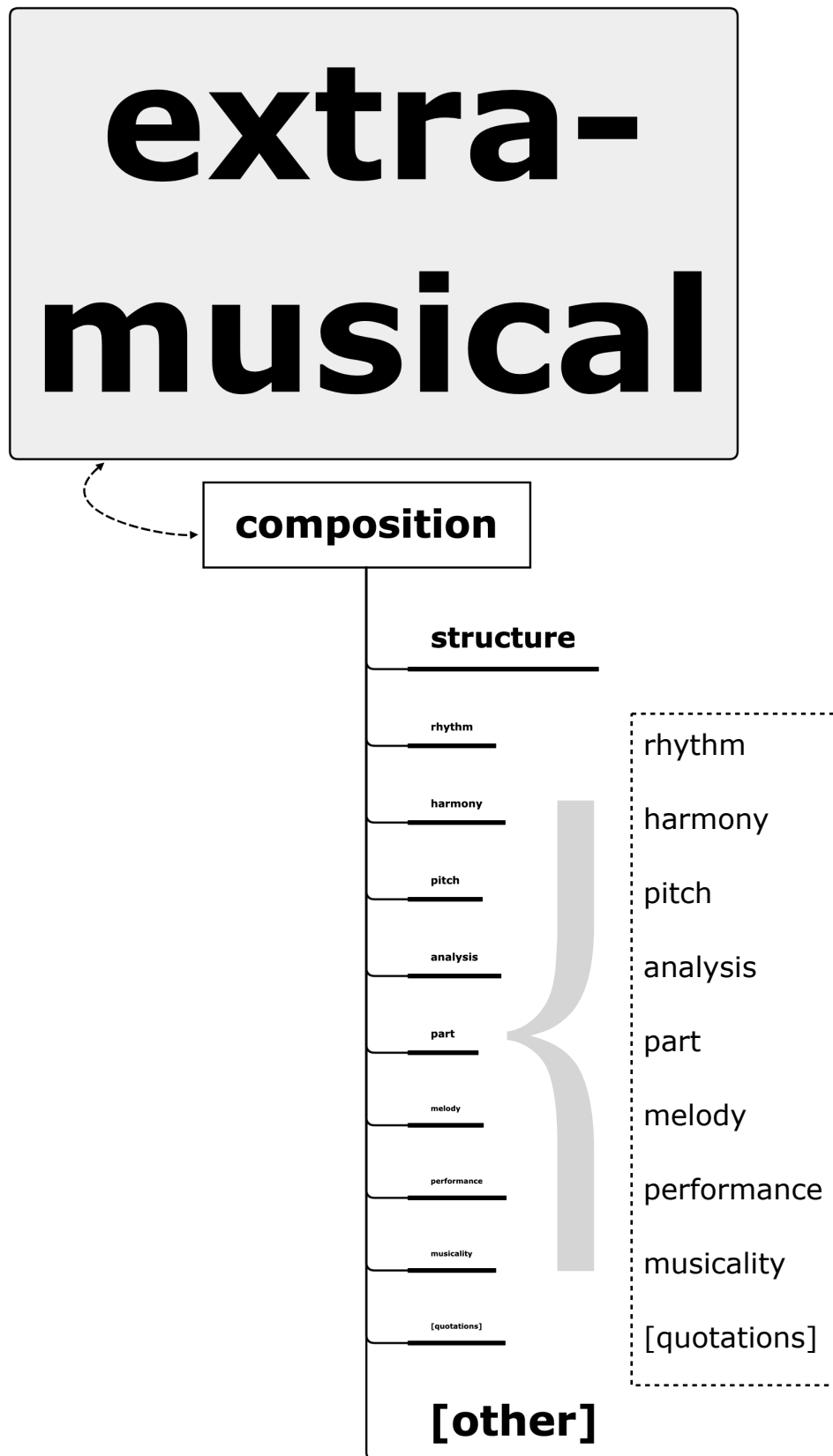
of its use as a reference to Medieval education, rather than to the inclusion of music in that part of a university education. Similarly, 'lament' is included in the 'composition' category because it references a type of piece. In cases where a word could be assigned to one of two or more categories, such as 'cadence' (which could be assigned to 'rhythm', 'harmony', or 'structure'), it has been placed in the category that represents the context in which it is used in each given instance. Lastly, hyphenated items represent instances in which a Classifier+Nominal serves as a complete semantic unit, the hyphenation being a means of retaining the unit during data processing.

There are a few noteworthy points to observe among the 48 cartographic categories in Table 4.10.2 above. Firstly, the sheer magnitude of the 'extramusical' category is somewhat misleading. While it is true that it includes a great deal of non-disciplinary specific content, it also contains numerous generic nouns. Thus, it cannot be claimed that all of this is contextual information. Nonetheless, it is conspicuously out of proportion to the next largest categories, composition and structure. This suggests that much of what is being written frequently (since this content is linked to frequent LB in the most prestigious and oft-cited journals) has little to do with the actual phenomenon of music, but rather with other

Table 4.10.3: Samples of Nominal Ambient Content by Category

Rank	Category	Samples
1	Extramusical	abbreviations, calendar, ease, galleries, I, kind, machine, object, quadrivium, sadness, U.S., wake, zeitgeist
2	Composition	ballad, dance-forms, fabula, hymn, lament, nocturne, parody, recitative, techno-dance, version
3	Structure	amen, cadence, ending, ground, interpolation, medleys, opening, s-period, unit
4	Rhythm	bar, demisemiquavers, half, long-short-short, note-values, pause, rest, talea, value
5	Harmony	anticipation, c-major, enharmonic-equivalence, key, major-mode, sequences, upper-neighbor
6	Pitch	bass-leaps, degree, f [forte], hexachord, note, pedal-pitch, row, tetrachord
7	Analysis	analysability, charge, entrance, gesture, idea, kopfton, material, sonata-theory, urlinie
8	Part	bass-line, line, part, tenor, viola-parts
9	Melody	ascent, cantus, elaboration, genera, ideé, maqamat, oboe-call
10	Performance	diction, fiddling, performance, recitation, virtuosity
11	Musicality	groove, music, sonorous

phenomenon being related to music through the Discourse itself. In other words, Music Discourse has rather less to do with the phenomenon of music than might be assumed on the strength of its name and position within academia. (*Note: Though anecdotal, this finding supports my personal experience, both of studying

Musicology and finding it rather disconnected from music-as-phenomenon, and the experience of performer friends who have said repeatedly that Musicology has nothing to do with music.) Of course, this raises an obvious question: If not largely about the phenomenon of music, then what is the focus of the Discourse? Though the answer to that lies beyond the scope of the present study, the occurrence of *century* bundles does hint that Music Discourse is highly concerned with history and therefore potentially with the development of musical ideas over time.

The second largest category, 'composition', encompasses both general references to works or pieces and all specific titles that appear in the ExCo. At 11% of the total, this is a significant category. This partially redresses the seeming lack of focus on the phenomenon of music posed by the weight of the 'extramusical' category; nevertheless, discussion of works, particularly titled works, once again indicates dependence on score notation. That is, the Discourse is focused on scores-as-music rather than performing-as-music or listening-as-music. That view privileges certain agents and aspects of music making, elevating them to prestigious positions within the Discourse, most especially composers. Here again is evidence that the Discourse remains centered on the Western canon.

The earlier observation that structure is a central concern of Music Discourse, particularly given the status afforded lengthy notated pieces (i.e., work, opera, symphony, sonata), is again confirmed by its prominence in the ambient content. The category of 'structure' includes sections, phrases, and various units, as shown in Word Cloud 4.10.1, which presents the highest-frequency terms from this category. All of the items in this category can be subjected to analysis, particularly harmonic analysis. Indeed, this is verified by the next five most prominent categories, each of which constitutes 3% of the total ambient content. In order of appearance, these are: rhythm, harmony, pitch, analysis, part. At last, rhythm makes an appearance at this level of analysis, along with three other items that constitute aspects of music analysis. Harmony again forms a substantial part of the Discourse as ambient content, since it appears at this level along with 'pitch' and 'part', both potential aspects of harmony, though also of melody. In fact, the final four categories, each of which accounts for 2% of the ambient content, are: melody, performance, musicality, and quotations. It is remarkable that quotations rival something as fundamental as melody for space in the Discourse. It is similarly striking that performance resides at this level, given that performers and audiences are critical agents in the making and receiving of music. If a piece is performed and no one hears it, is it music?

Having considered the categories that are most prominent, it is also instructive to look at those which are least so. Returning to Table 4.10.2 above, categories that account for less than 0.1% are, in order: critic, organization, trend, producer, transmission, author, community, composer/performer, editor, manager, tuning, recordist. There are several agents here who are involved in the process of producing, transmitting, and receiving musical performances, yet they are barely mentioned.

Among the 24 categories that fall between the most frequent 12 and least frequent 12 (i.e., those mentioned in the preceding paragraph), it is intriguing to note that overt references to emotion occur a mere 29 times. Given its ability to provoke strong feelings, it seems curious that more mention is not made of the effect music has on listeners. Also in this middle 50% of categories are other aspects of music that have been absent until this stage of analysis, including several additional agents in the music making process (e.g., performers, composers, listeners, musicians), instruments and their performance techniques, and various fundamental facets of music as a phenomenon (i.e., dynamics, texture, tempo, ornamentation, articulation, timbre).

Word Cloud 4.10.1: High-Frequency Terms in 'Structure' Category

https://worditout.com/word-cloud/5510065

The relationships between all 48 categories is visualized in Map 4.10.1 as an interplay between the agents and means of producing music, technologies for transmission of it, and participants and aspects of its reception. Five magnifications (i.e., Map 4.10.1 Magnifications 1-5) follow the full map to permit easier examination of its details. Though notation is mapped under Transmission, as a means of permanently preserving composition, it could also have been tethered to composition with all its elements that are notated in scores. Similarly, the various items grouped together in the red

dotted-outline box attached to sound could be listed as elements of music. They are not listed as such here, however, on account of their infrequent appearance in the Discourse. As rare items, they appear not to be considered primary elements of music. Thus, I have chosen to associate them generally with sound and performance, respectively.

Map 4.10.1 Cartography of Ambient Content

Composition

Production

MUSIC

Elements

Sound

Extra-musical

Structure

Performance

Instrumentation

Instrument

Timbre

Texture

Song

Perf

Musicality

Performer

Notation

Ornamentation

Dynamic

Articulation

Technique

Timbre

Tuning

Text

Community

Author

Style

Melody

Harmony

Composer

Rhythm

Pitch

Analysis

Musicologist

Notation

Emotion

Trend

Listener

Criticism

Critic

Transmission

Organization

Manager

Whole

Recording

Notation

Reporter

Producer

Editor

Map 4.10.1 Magnification 1: Reception
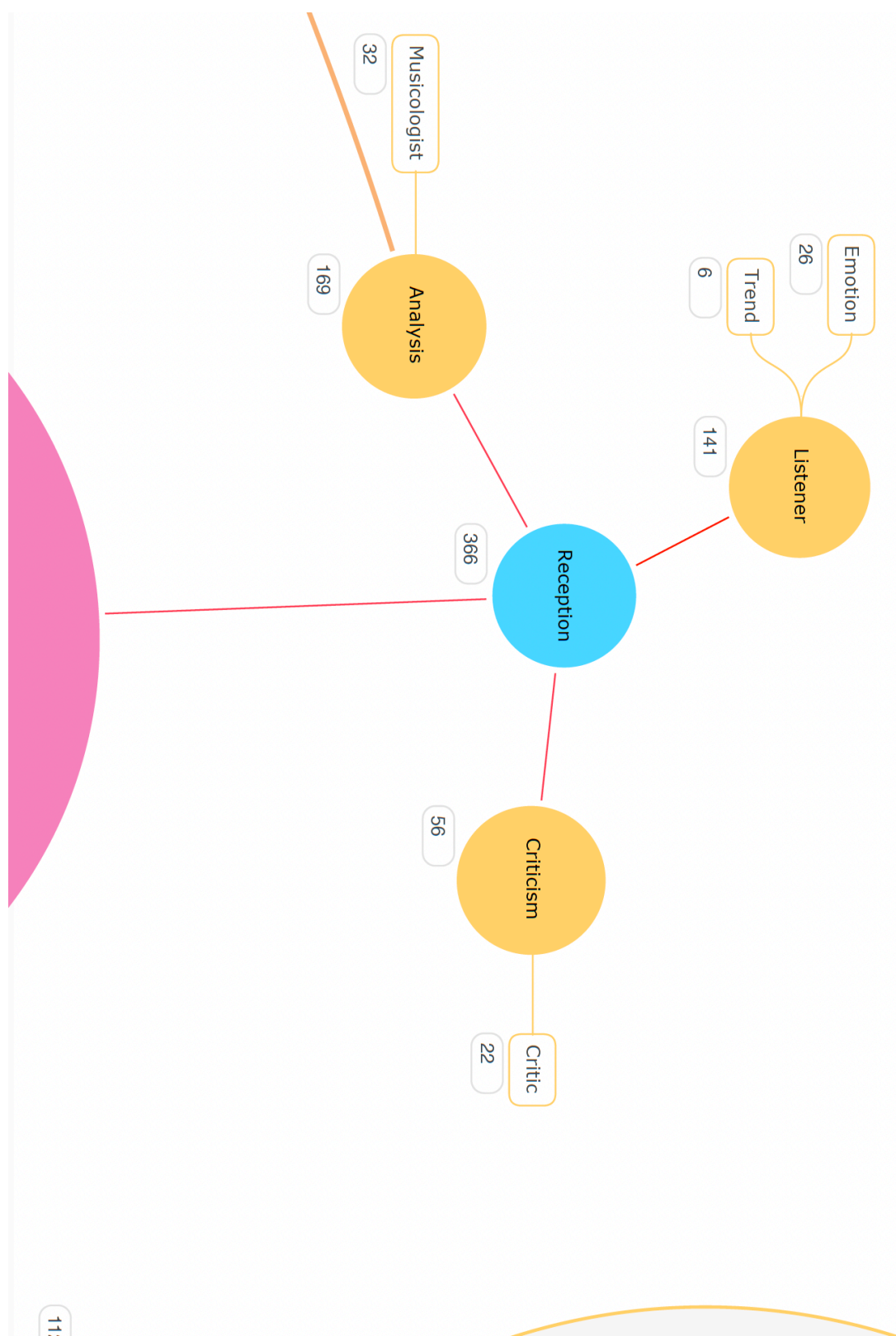
Map 4.10.1 Magnification 2: Sound
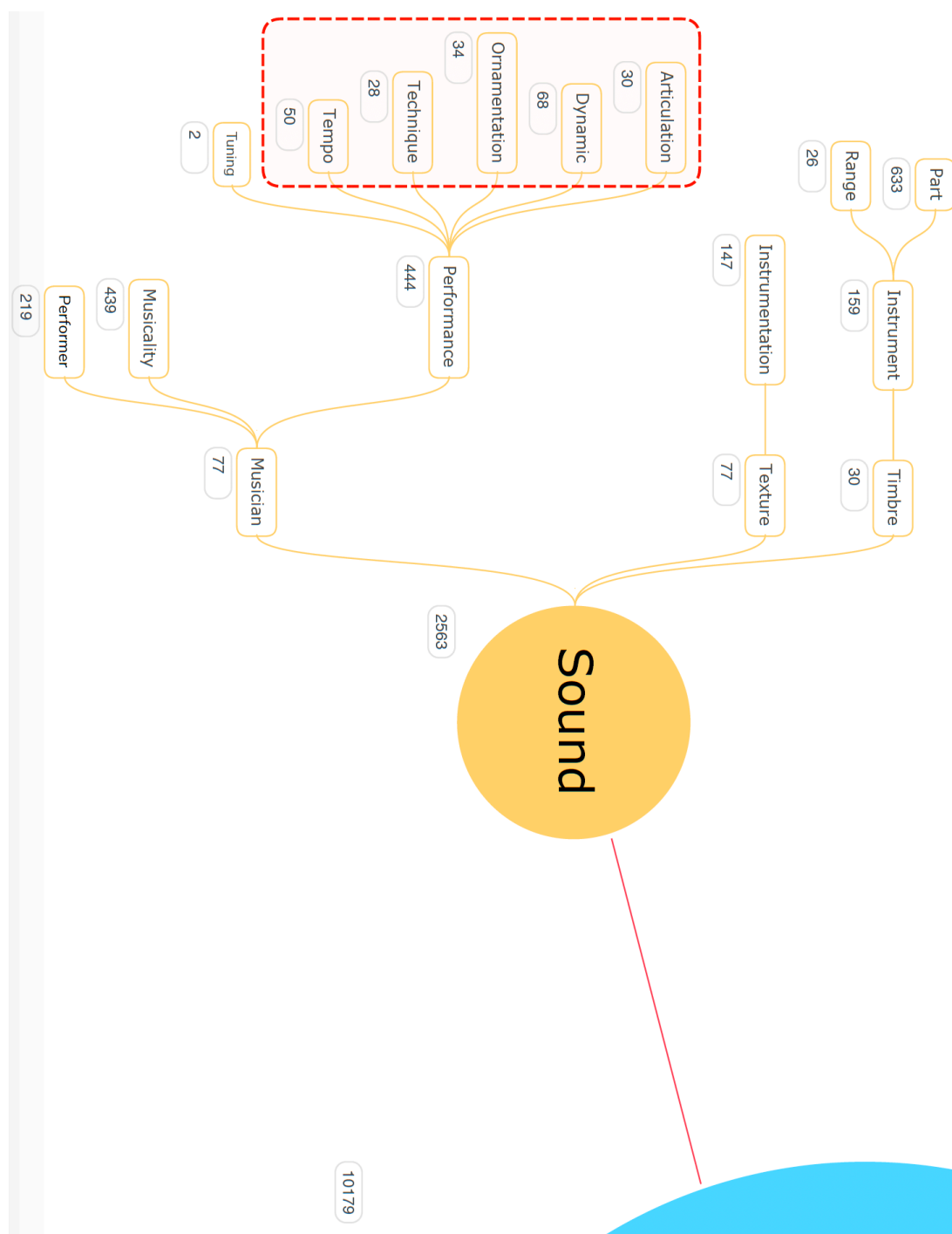
Map 4.10.1 Magnification 3: Composition

3026

Composition

Style 199

2 Community

130 Text

Author 2
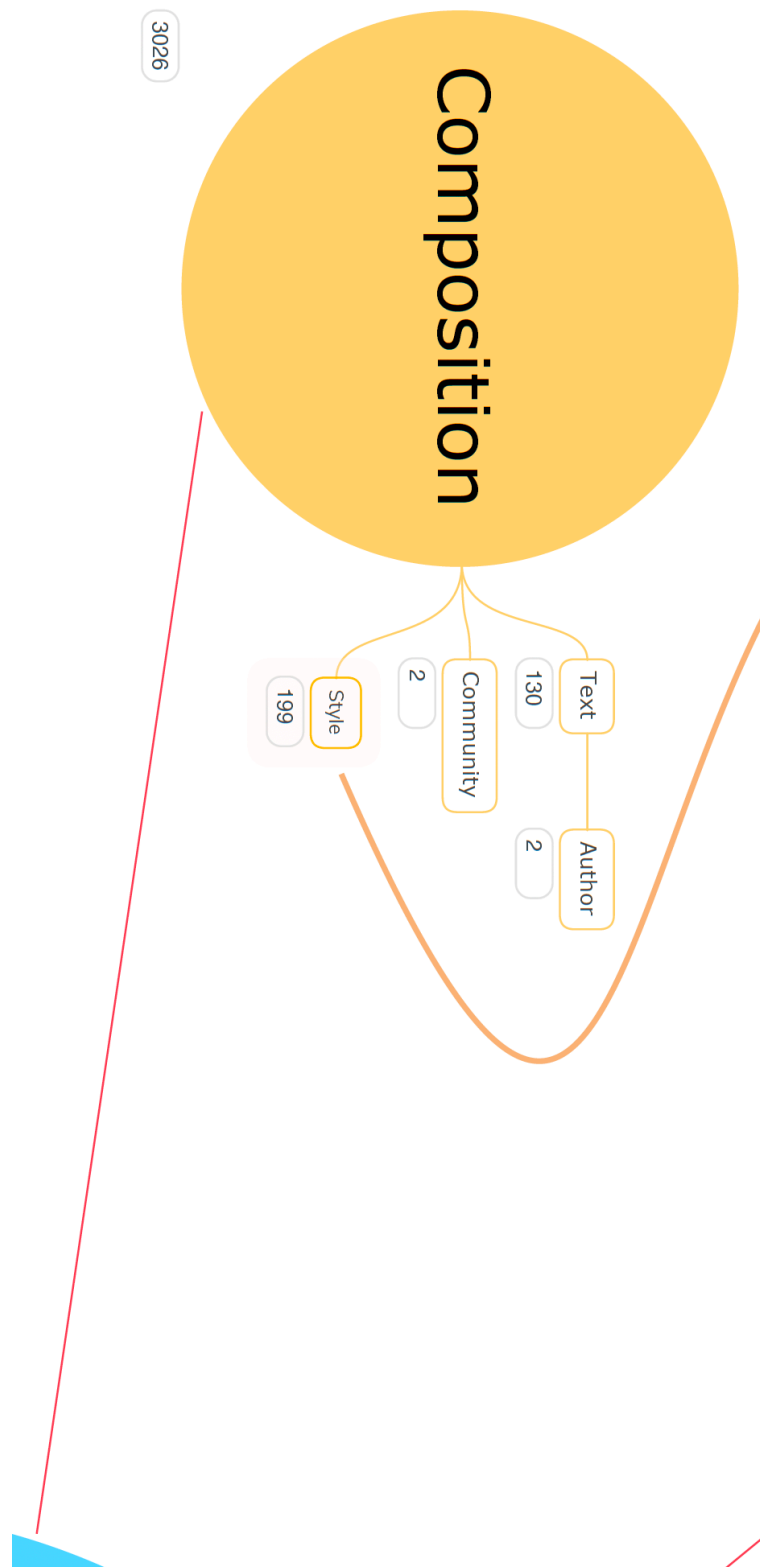
Map 4.10.1 Magnification 4: Elements

1848 — Structure

98 — Counterpoint
754 — Harmony
584 — Melody

738 — Pitch
666 — Rhythm

4590

Elements
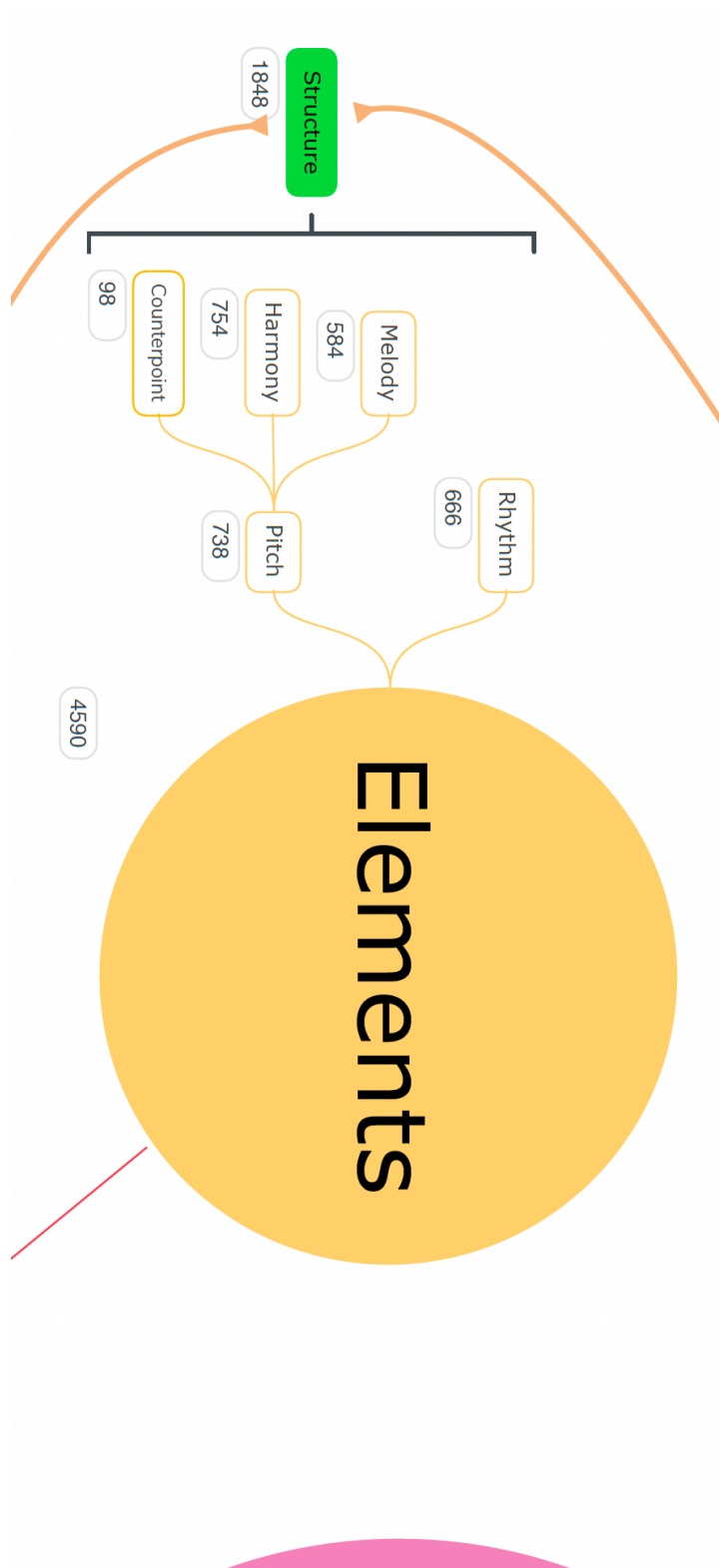
Map 4.10.1 Magnification 5: Transmission
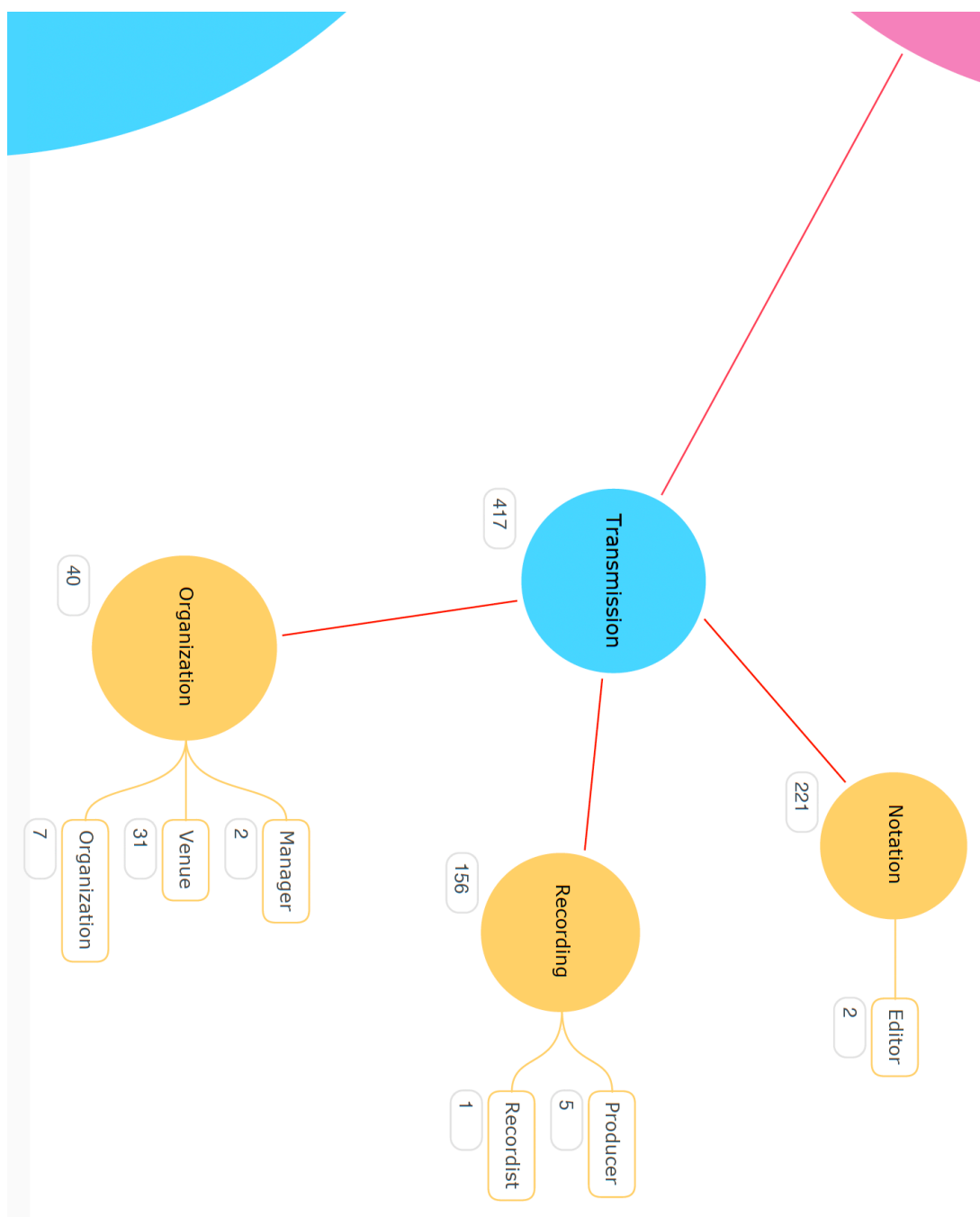
4.11 Summary of Expert Corpus Findings

Having thoroughly reviewed the findings from the Expert Corpus, it is now possible to present a succinct summary of Music Discourse. Most notably, these findings point unequivocally to the fact that it is delimited by its reliance on the semiotic of score notation. Non-disciplinary lexical bundles substantially outnumber their disciplinary counterparts largely because the former typically represent parts of functional word groups or complete groups. Such is the case with bundles that are composed of prepositions and deictics, or complete verbal groups. Among bundle structural Types, there is also a high proportion of genitive and potential genitive bundles. Of those that include generic nominals, an unexpectedly high proportion contain numeratives. Indeed, even these numerative bundles outnumber disciplinary ones. This clearly indicates their centrality to the Discourse, as reflected by the fact that all 5-LB function as numeratives as do a significant majority of the extended 6- to 8-LB. Among these, by far the largest percentage are portion bundles, mostly through measurement of portions of scores. In addition to portion bundles, there exist a group of *century* bundles that demonstrate not only a teleological view of music but that also largely confine the Discourse to a limited span of time, from the Eighteenth to the early Twentieth Centuries, a fact that demonstrates a fascination with tonal harmony and the so-called

canon of Western Music. This view is well-supported by the content of the disciplinary bundles, the majority of which are concentrated on form and harmony. This conclusion is further supported by the dominance of masculine pronouns and American and British geographic locations within the LB. Global views of complete texts reveal that numerative and disciplinary bundles increase significantly in density from Ethnomusicology articles, to Musicology, to Music Theory, which is another clear demonstration of the centrality of score notation in the Discourse, given that the music studied by Ethnomusicologists is typically not notated in staff notation. This also indicates a hierarchy in which the study of music notated thusly is afforded greater prestige. The final confirmation that Music Discourse is delimited by score notation comes from analysis of ambient content. The vast majority of bundles are referential in nature, connecting to propositional content in the slots immediately preceding and succeeding bundles. Out of 22,711 nominals drawn from those slots, 65% are filled with extramusical content, compositions, or structural terms, respectively. The last of these are comprised of various portions of musical works or pieces, thus corroborating both the importance and function of numeratives in the Discourse.

Given all of the above, it is now possible to state not only that large portions of Music Discourse are not about music, but that those that are offer, at most, a highly delimited view of music, one that is dependent on Western score notation. Thus, the music created with such notation is clearly perceived as prestigious by expert writers, statements and protestations to the contrary notwithstanding. What the overall focus of the Discourse is remains somewhat unclear, as an exhaustive examination of the extramusical ambient content lies beyond the purview of this study. Given that so many fundamental facets of music are not prioritized, however, there is an overwhelmingly narrow focus on structure and harmony; a fact corroborated by the presence of multiple *century* bundles (nineteenth, twentieth, and eighteenth centuries, respectively), strongly suggesting a narrow focus on music of the tonal and post-tonal periods. Indeed, further instances of these bundles are found among the ambient content, as shown in Table 4.11.1, which yet again displays a predilection for music of the Eighteenth to Twentieth Centuries, and therefore the tonal harmony of that period. While harmony can be evocative and structure can facilitate coherence for the listener, all the other facets of music that are left largely unspoken (even ones that can be, and often are notated in scores, such as dynamics, tempo, ornamentation, articulation and timbre) are vital to composers, performers, and listeners alike. By

largely ignoring them, much of what differentiates musical pieces and makes them attractive is ignored. Thus, dependence on score notation for the purpose of examining harmonic form severely restricts the scope of Music Discourse and consequently the view of music it presents to the scholarly community.

Though there are items herein that may well warrant functional analysis at a more delicate level, this has been left aside precisely because LB as frequency items lack context, which renders further analysis difficult if not moot for the present study.

Table 4.11.1: Century Bundles among Ambient Content

Frequency	Century	Frequency	Century
76	20th century	7	17th century
71	19th century	5	21st century
13	18th century	3	14th century
7	15th century	2	12th century
7	16th century	2	13th century

5 Comparison of ExCo and NoCo Findings

5.1 Introduction

Though the Novice Corpus is considerably smaller than the Expert Corpus, there is still sufficient data therein to sustain a lengthy discussion of its findings, as well as to support conclusions regarding comparisons with the ExCo. These comparisons serve two purposes. The first is to evaluate the degree to which novice writing approximates expert writing, which is a measure of the success of both expert teaching and novice learning. Since the production of expert text — in this case, disciplinary writing — serves as a gate-keeping function within a given discourse community (Gee 2014a), it is critical to assess the application of this function to the learning and acquisition process. The second purpose is to evaluate how well L2 novices understand the practices of the discourse they are asked to learn, specifically the interaction of lexical bundles with propositional content.

5.2 Lexical Bundle Type, Frequency, Range

The Novus Corpus (NoCo) is composed of eighteen texts: eight from secondary -level students, ten from tertiary-level students. The lowest feasible minimums were set: a Frequency of 4 and a Range of 3. This resulted in 166 Types with 1,174 Tokens. As several students in each group wrote in response to a common prompt,

however, all LB drawn exclusively from either the secondary or tertiary sub-corpora (87 in total) were eliminated so as to prevent a single prompt from unduly influencing the Corpus returns. Additionally, 14 other LB were eliminated because they are either bisected by punctuation, which AntConc does not recognize, or because they are subsumed under longer bundles (N.B.: One Token of NoCo 33 is not subsumed but is part of a quotation and was therefore eliminated). Table 5.2.1 shows the distribution of the remaining 65 LB by number of words per bundle.

Table 5.2.1: Number of LB Types by words per bundle

	3-WORD	4-WORD	5-WORD
No. of Types	56	8	1

All of the above elimination procedures, intended to ensure the representativeness of the NoCo, appear justified by the resulting proportion of 3-, to 4-, to 5-LB: 56:8:1, as visualized in Chart 5.2.1. As with the ExCo LB, these exhibit nearly the same proportions, increasing approximately by an order of magnitude for each successively shorter bundle.

Table 5.2.2 lists all 65 remaining LB Types with their frequencies and ranges. The most frequent item, NoCo 1 'played by the', accounts for 5% of all 515 Tokens, though only across a range of seven texts,

269 of 460

Chart 5.2.1: Proportion of n-LB (65 Types)

2% 12%

86%

● 5-LB ● 4-LB ● 3-LB
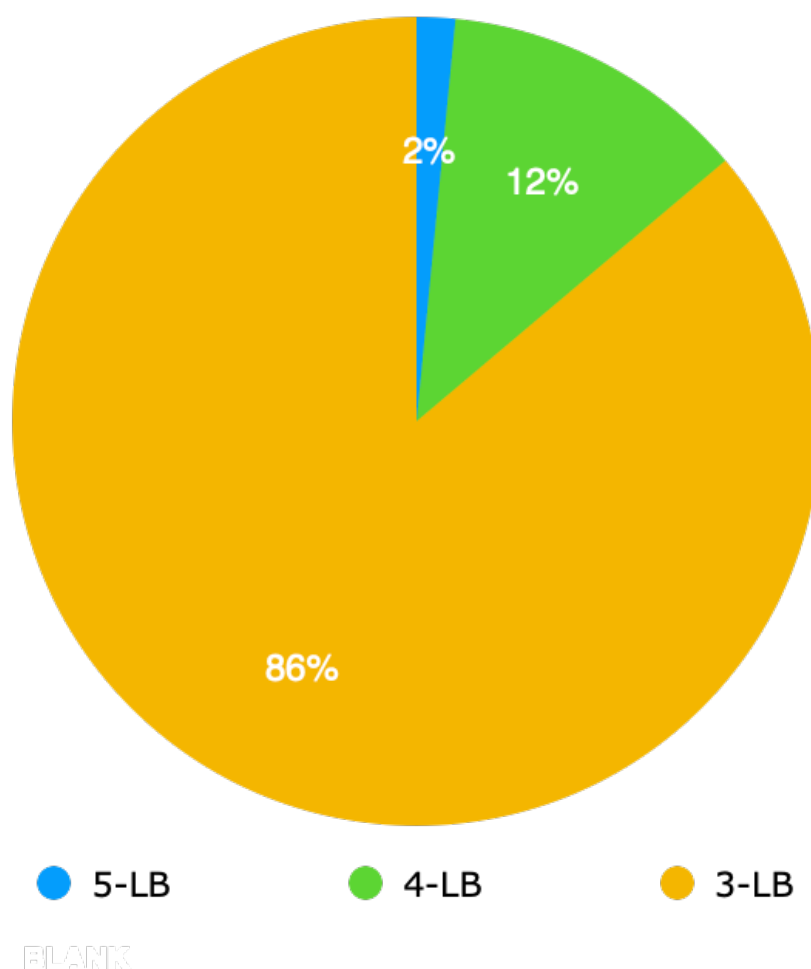
which suggests an excessive influence from one sub-corpus. This proves true, as just one of those Tokens comes from the tertiary sub-corpus. Nevertheless, this proves instructive as it hints that there may be differing degrees of engagement with aspects of music, such as performance, by educational level. As demonstrated in the previous chapter, the actual performance of music is not a significant dimension of Expert Discourse. The frequency of this bundle may also suggest that writing at the secondary level is intended to engage novices with music they are learning to perform.

That said, this is the only obviously disciplinary-specific LB in the most frequent 10% of NoCo bundles.

Following this first LB with its unexpectedly high frequency, the next several high-frequency Types are all predictable, with several participating in extended numeratives already familiar from the ExCo. Additionally, extended numeratives appear across the largest range of texts. To confirm whether these high-frequency correlations between NoCo and ExCo represent an alignment of novice writing to expert writing, it is necessary to compare which LB appear in both Corpora and in what order.

Table 5.2.3 compares the NoCo and ExCo LB, demonstrating which bundles are shared. Of the 65 NoCo bundles, 39 are shared by the ExCo, a full 60%. These include the single 5-LB (highlighted in pink), half of the eight 4-LB (highlighted in orange), and 34 of the 56 3-LB. Among the most frequent bundles that are shared, there is a relatively high correlation of rank within both Corpora, many of which are numeratives. Such a high degree of correspondence already suggests that these novices have acquired several high-frequency bundles from the Discourse. While this is predictable, given that lexical bundles are likely memorized and retrieved as units (Wood 2015), it is noteworthy that these high-frequency items

Table 5.2.2: NoCo LB by Frequency and Range

Rank	Fre-quency	Range	S1	S2	S3	S4	S5
1	27	7	play-ed	by	the		
4	19	12	the	end	of		
8	16	8	in	the	first		
10	16	8	the	same	time		
12	14	8	at	the	same	time	
13	14	8	one	of	the		
14	14	4	the	first	move-ment		
15	14	10	the	use	of		
16	13	7	a	lot	of		
17	13	11	at	the	end		
18	13	7	end	of	the		
22	12	7	the	end	of	the	
25	11	6	the	second	move-ment		
26	11	5	the	third	move-ment		
27	11	3	there	is	a		
28	10	6	is	played	by		
29	10	6	the	begin-ning	of		
32	9	4	and	it	is		
34	9	9	at	the	end	of	
37	9	6	in	the	second		
38	9	5	in	the	third		
39	9	7	of	the	first		
40	9	7	of	the	music		

Table 5.2.2: NoCo LB by Frequency and Range

Rank	Fre-quency	Range	S1	S2	S3	S4	S5
41	9	5	of	the	piece		
44	8	5	at	the	beginning	of	
46	8	4	by	the	violin		
49	8	5	in	bars	[#]	and	
55	7	4	can	be	seen		
57	7	4	in	order	to		
60	7	4	the	second	theme		
62	7	4	the	theme	and		
63	7	4	the	violin	and		
64	6	4	a	group	of		
69	6	4	be-tween	the	two		
70	6	4	con-trast	with	the		
72	6	5	from	the	first		
74	6	3	play-ed	by	the	violin	
82	6	5	with	the	first		
83	5	3	and	this	is		
84	5	4	as	well	as		
86	5	3	follow-ed	by	a		
89	5	4	in	the	second	move-ment	
90	5	3	in	the	third	move-ment	
92	5	4	is	first	intro-duced		
93	5	4	is	playing	the		

Table 5.2.2: NoCo LB by Frequency and Range

Rank	Fre- quency	Range	S1	S2	S3	S4	S5
94	5	4	it	is	the		
101	5	3	the	sound	of		
106	5	3	the	violin	in		
108	5	5	use	of	the		
109	4	3	A	in	the		
113	4	3	and	so	on		
114	4	4	and	then	the		
116	4	4	as	the	first		
117	4	3	at	the	beginning	of	the
127	4	3	in	the	introduc- tion		
132	4	4	is	one	of		
135	4	4	it	is	a		
139	4	4	of	the	original		
140	4	4	of	the	second		
143	4	3	of	this	move- ment		
147	4	3	simi- lar	to	the		
153	4	3	the	idea	of		
155	4	3	the	lower	strings		
161	4	4	there	is	also		
164	4	3	to	deve- lop	the		

do not seem to be overused, as they tend to appear twice per text, occasionally thrice.

Table 5.2.3: NoCo and ExCo Shared LB with Frequency and Range (% of Texts)

NoCo				LB	ExCo			
Rank	Frequency	Range	% Texts		Rank	Frequency	Range	% Texts
4	19	12	67%	the end of	3	439	119	74%
8	16	8	44%	in the first	5	344	113	71%
10	16	8	44%	the same time	20	216	96	60%
12	14	8	44%	at the same time	25	206	95	59%
13	14	8	44%	one of the	2	527	139	87%
14	14	4	22%	the first movement	56	137	31	19%
15	14	10	56%	the use of	27	202	90	56%
17	13	11	61%	at the end	11	263	88	55%
18	13	7	39%	end of the	8	294	107	67%
22	12	7	39%	the end of the	13	259	99	62%
27	11	3	17%	there is a	35	173	72	45%
29	10	6	33%	the beginning of	16	226	89	56%
32	9	4	22%	and it is	164	82	62	39%
34	9	9	50%	at the end of	15	226	85	53%
37	9	6	33%	in the second	28	201	82	51%
38	9	5	28%	in the third	254	64	37	23%
39	9	7	39%	of the first	9	284	101	63%

Table 5.2.3: NoCo and ExCo Shared LB with Frequency and Range (% of Texts)

NoCo				LB	ExCo			
Rank	Frequency	Range	% Texts		Rank	Frequency	Range	% Texts
40	9	7	39%	of the music	10	270	98	61%
41	9	5	28%	of the piece	24	209	52	33%
44	8	5	28%	at the beginning of	74	125	61	38%
55	7	4	22%	can be seen	189	73	44	28%
57	7	4	22%	in order to	7	308	117	73%
64	6	4	22%	a group of	243	65	36	23%
69	6	4	22%	between the two	88	114	67	42%
72	6	5	28%	from the first	392	49	30	19%
82	6	5	28%	with the first	480	44	30	19%
84	5	4	22%	as well as	1	537	140	88%
86	5	3	17%	followed by a	127	94	47	29%
94	5	4	22%	it is the	67	130	70	44%
101	5	3	17%	the sound of	219	69	38	24%
108	5	5	28%	use of the	147	87	56	35%
113	4	3	17%	and so on	349	53	37	23%
117	4	3	17%	at the beginning of the	239	66	42	26%

Table 5.2.3: NoCo and ExCo Shared LB with Frequency and Range (% of Texts)

NoCo				LB	ExCo			
Rank	Frequency	Range	% Texts		Rank	Frequency	Range	% Texts
132	4	4	22%	is one of	110	103	63	39%
135	4	4	22%	it is a	79	121	69	43%
139	4	4	22%	of the original	130	93	51	32%
140	4	4	22%	of the second	61	135	60	38%
147	4	3	17%	similar to the	407	48	29	18%
153	4	3	17%	the idea of	38	161	76	48%

One other point of interest based on this view is the relative correspondence of the range percentages, which give the percent of the total number of Corpus texts each Type of LB represents. Only 12 of the total shared bundles (30%) have a discrepancy greater than 20%, with many having percentages in close proximity, again suggesting a degree of competence on the part of novices. This supports Wray's (2012) assertion that lexical bundles are key to surviving academia.

5.3 Word Class Analysis

Table 5.3.1 presents a word class analysis of the NoCo LB (rank numbers of LB shared with the ExCo are highlighted in blue). This is followed by Table 5.3.2, which displays the percentage of word classes by slot for 5-LB. Table 5.3.3 shows the single syntagmatic Type of the 5-LB. Because this bundle is also present in the ExCo, there is little to add here regarding its structure.

The same is partially true for the word class (Table 5.3.4) and Syntagmatic Type (Table 5.3.5) analyses of the 4-LB, as four of them are shared in common with the ExCo. Because the remaining four are not, however, a brief discussion of the 4-LB Syntagmatic Types is warranted. The eight LB Tokens form six distinct syntagmatic Types (S-Types), all of which contain a noun and two of which contain numbers. Of the six S-Types, four are shared with the ExCo. The remaining two, S-Types 4.5 and 4.6, represent two phenomena unique to the NoCo: 1) retention of a LB that includes variable bar numbers in a score (NoCo 49 'in bars [#] and'), and 2) a LB referencing performance (74 'played by the violin'). The former has been retained because it is an example of how novices learn to reference measure numbers, and thus it potentially offers a brief glimpse of writing instructor's priorities.

Table 5.3.1: NoCo LB Word Class Analysis

RANK	Lexical Bundle					Word Class				
	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5
1	play-ed	by	the			V	Prep	Det		
4	the	end	of			Det	N	Prep		
8	in	the	first			Prep	Det	Num		
10	the	same	time			Det	Adj	N		
12	at	the	same	time		Prep	Det	Adj	N	
13	one	of	the			Num	Prep	Det		
14	the	first	move-ment			Det	Num	N		
15	the	use	of			Det	N	Prep		
16	a	lot	of			Det	N	Prep		
17	at	the	end			Prep	Det	N		
18	end	of	the			N	Prep	Det		
22	the	end	of	the		Det	N	Prep	Det	
25	the	se-cond	move-ment			Det	Num	N		
26	the	third	move-ment			Det	Num	N		
27	there	is	a			Prn	V	Det		
28	is	play-ed	by			V	V	Prep		
29	the	begin-ning	of			Det	N	Prep		
32	and	it	is			Conj	Prn	V		
34	at	the	end	of		Prep	Det	N	Prep	
37	in	the	se-cond			Prep	Det	Num		

Table 5.3.1: NoCo LB Word Class Analysis

RANK	Lexical Bundle					Word Class				
	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5
38	in	the	third			Prep	Det	Num		
39	of	the	first			Prep	Det	Num		
40	of	the	music			Prep	Det	N		
41	of	the	piece			Prep	Det	N		
44	at	the	begin-ning	of		Prep	Det	N	Prep	
46	by	the	violin			Prep	Det	N		
49	in	bars	[#]	and		Prep	N	Num	Conj	
55	can	be	seen			V	V	V		
57	in	order	to			Prep	N	Prep		
60	the	se-cond	theme			Det	Num	N		
62	the	theme	and			Det	N	Conj		
63	the	violin	and			Det	N	Conj		
64	a	group	of			Det	N	Prep		
69	be-tween	the	two			Prep	Det	Num		
70	con-trast	with	the			N	Prep	Det		
72	from	the	first			Prep	Det	Num		
74	play-ed	by	the	vio-lin		V	Prep	Det	N	
82	with	the	first			Prep	Det	Num		
83	and	this	is			Conj	Det	V		
84	as	well	as			Prep	Adv	Prep		
86	fol-low-ed	by	a			V	Prep	Det		

Table 5.3.1: NoCo LB Word Class Analysis

R A N K	Lexical Bundle					Word Class				
	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5
89	in	the	se-cond	mov-e-ment		Prep	Det	Num	N	
90	in	the	third	mov-e-ment		Prep	Det	Num	N	
92	is	first	intro-duced			V	Num	V		
93	is	play-ing	the			V	V	Det		
94	it	is	the			Prn	V	Det		
101	the	sound	of			Det	N	Prep		
106	the	violin	in			Det	N	Prep		
108	use	of	the			N	Prep	Det		
109	A	in	the			N	Prep	Det		
113	and	so	on			Conj	Det	Prep		
114	and	then	the			Conj	Adv	Det		
116	as	the	first			Prep	Det	Num		
117	at	the	begin-ning	of	the	Prep	Det	N	Pre p	De t
127	in	the	intro-duction			Prep	Det	N		
132	is	one	of			V	Num	Prep		
135	it	is	a			Prn	V	Det		
139	of	the	origi-nal			Prep	Det	Adj/ N		
140	of	the	se-cond			Prep	Det	Num		
143	of	this	move-ment			Prep	Det	N		
147	simi-lar	to	the			Adj	Prep	Det		

Table 5.3.1: NoCo LB Word Class Analysis

RANK	Lexical Bundle					Word Class				
	S1	S2	S3	S4	S5	S1	S2	S3	S4	S5
153	the	idea	of			Det	N	Prep		
155	the	lower	strings			Det	Adj	N		
161	there	is	also			Det	V	Adv		
164	to	develop	the			V infin	V infin	Det		

Turning to the 3-LB, the word class and syntagmatic analyses for these are presented in Table 5.3.6 and Table 5.3.6, respectively, for all 56 bundles. Unlike the

Table 5.3.2: NoCo 5-LB Word Classes by Slot (1 LB Type)

Class	Slot 1	Slot 2	Slot 3	Slot 4	Slot 5
Det		1			1
N			1		
Prep	1			1	

Table 5.3.3: NoCo 5-LB Syntagmatic Type (1 LB Type)

Slot 1	Slot 2	Slot 3	Slot 4	Slot 5
Prep	Det	N	Prep	Det

5-LB and 4-LB, there are numerous S-Types not shared with the ExCo. In fact, 12 of the total 26 Types (46%) are not shared. Among those, all but one is represented by a single LB. The

Table 5.3.4: NoCo 4-LB Word Classes by Slot (8 LB Types)

Class	Slot 1 #	Slot 1 % of 8	Slot 2 #	Slot 2 % of 8	Slot 3 #	Slot 3 % of 8	Slot 4 #	Slot 4 % of 8
Adj					1	12.5%		
Conj							1	12.5%
Det	1	12.5%	5	62.5%	1	12.5%	1	12.5%
N			2	25%	2	25%	4	50%
Num					3	37.5%		
Prep	6	75%	1	12.5%	1	12.5%	2	25%
V	1	12.5%						

remaining one is represented by two LB (S-Type 3.7: determiner, noun, conjunction). The fact that these 12 are not shared with the ExCo and that each Type is scarce strongly suggests that they are unique to the NoCo because they represent a degree of inexperience with more complex lexicogrammatical constructions. This conjecture is supported by Type 3.7 with its single noun followed by a conjunction since both examples contain simple nominal parataxis.

Table 5.3.5: NoCo 4-LB Syntagmatic Types (8 LB Types)

No.	#	% of 8	Slot 1	Slot 2	Slot 3	Slot 4
4.1	1	12.5%	Det	N	Prep	Det
4.2	1	12.5%	Prep	Det	Adj	N
4.3	2	25%	Prep	Det	N	Prep
4.4	2	25%	Prep	Det	Num	N
4.5	1	12.5%	Prep	N	Num	Conj
4.6	1	12.5%	V	Prep	Det	N

The two LB constituting that Type are NoCo 62 'the theme and' and 63 'the violin and'. This analysis, then, offers the first evidence for a distinctly novice approach to writing, one that may not meet expert standards.

Table 5.3.6: NoCo 3-LB Word Classes by Slot (56 LB Types)

Class	Slot 1		Slot 2		Slot 3	
	#	% of 56	#	% of 56	#	% of 56
Adj	1	2%	2	4%		
Adj/N					1	2%
Adv			2	4%	1	2%
Conj	4	7%			2	4%
Det	17	30%	18	32%	14	25%
N	4	7%	12	21%	12	21%
Num	1	2%	5	9%	9	16%
Prep	18	32%	8	14%	13	23%
Prn	3	5%	1	2%		
V	8	14%	8	14%	4	7%

Table 5.3.7: NoCo 3-LB Syntagmatic Types (56 LB Types)

No.	#	% of 56	Slot 1	Slot 2	Slot 3
3.1	1	>2%	Adj	Prep	Det
3.2	1	>2%	Conj	Adv	Det
3.3	1	>2%	Conj	Det	Prep
3.4	1	>2%	Conj	Det	V
3.5	1	>2%	Conj	Prn	V
3.6	2	4%	Det	Adj	N
3.7	2	4%	Det	N	Conj
3.8	8	14%	Det	N	Prep
3.9	4	7%	Det	Num	N
3.10	1	>2%	Det	V	Adv
3.11	4	7%	N	Prep	Det
3.12	1	>2%	Num	Prep	Det
3.13	6	11%	Prep	Det	N
3.14	9	16%	Prep	Det	Num
3.15	1	>2%	Prep	Adv	Prep
3.16	1	>2%	Prep	Det	Adj/N
3.17	1	>2%	Prep	N	Prep
3.18	3	5%	Prn	V	Det
3.19	1	>2%	V	Num	Prep
3.20	1	>2%	V	Num	V
3.21	2	4%	V	Prep	Det
3.22	1	>2%	V	V	Prep
3.23	1	>2%	V	V	V
3.24	1	>2%	V	V	Prep
3.25	1	>2%	V	V	Det
3.26	1	>2%	V infin	V infin	Det

5.4 Word Function Analysis

Table 5.4.1 shows the functional analysis of words for all NoCo LB together. As with the analysis of the ExCo LB word functions, the two word functions examined most closely in the NoCo bundles are nominals and numeratives. These are marked in green, along with three verbals (each a form of 'to play'), to highlight disciplinary-specific content. The nominals and numeratives include a classifier (1), compound numeratives (6), ordinatives (19) and quantitatives (3), general nominals (8) and specific Things (12). As with the ExCo LB, numeratives figure prominently among the NoCo bundles, indicating one way in which novice writing aligns with expert, even to the extent of employing compound numeratives (ordinative + quantitative). The number of word functions per slot is listed in Table 5.4.2. Nominals are highlighted in green.

Table 5.4.1: NoCo Word Functions

RANK	Lexical Bundle	Function				
		S1	S2	S3	S4	S5
1	played by the	verbal	preposition	deictic		
4	the end of	deictic	ordinative	preposition		
8	in the first	preposition	deictic	ordinative		
10	the same time	deictic	post-deictic	nominal		

Table 5.4.1: NoCo Word Functions

RANK	Lexical Bundle	Function				
		S1	S2	S3	S4	S5
12	at the same time	preposition	deictic	post-deictic	nominal	
13	one of the	quantitative	preposition	deictic		
14	the first movement	deictic	compound numerative (ordinative quantitative)			
15	the use of	deictic	nominal	preposition		
16	a lot of	deictic	quantitative	preposition		
17	at the end	preposition	deictic	ordinative		
18	end of the	ordinative	preposition	deictic		
22	the end of the	deictic	ordinative	preposition	deictic	
25	the second movement	deictic	compound numerative (ordinative quantitative)			
26	the third movement	deictic	compound numerative (ordinative quantitative)			
27	there is a	deictic	verbal	deictic		
28	is played by	verbal group		preposition		
29	the beginning of	deictic	ordinative	preposition		
32	and it is	conjunction	deictic	Verbal		

Table 5.4.1: NoCo Word Functions

RANK	Lexical Bundle	Function				
		S1	S2	S3	S4	S5
34	at the end of	preposition	deictic	ordinative	preposition	
37	in the second	preposition	deictic	ordinative		
38	in the third	preposition	deictic	ordinative		
39	of the first	preposition	deictic	ordinative		
40	of the music	preposition	deictic	thing		
41	of the piece	preposition	deictic	thing		
44	at the beginning of	preposition	deictic	ordinative	preposition	
46	by the violin	preposition	deictic	thing		
49	in bars [#] and	preposition	thing	ordinative	conjunction	
55	can be seen	verbal group				
57	in order to	complex conjunction				
60	the second theme	deictic	compound numerative (ordinative quantitative)			
62	the theme and	deictic	thing	conjunction		
63	the violin and	deictic	thing	conjunction		
64	a group of	deictic	nominal	preposition		

Table 5.4.1: NoCo Word Functions

RANK	Lexical Bundle	Function				
		S1	S2	S3	S4	S5
69	between the two	preposition	deictic	quantitative		
70	contrast with the	nominal	preposition	deictic		
72	from the first	preposition	deictic	ordinative		
74	played by the violin	verbal	preposition	deictic	thing	
82	with the first	preposition	deictic	ordinative		
83	and this is	conjunction	deictic	verbal		
84	as well as	complex preposition				
86	followed by a	verbal	preposition	deictic		
89	in the second movement	preposition	deictic	compound numerative (ordinative quantitative)		
90	in the third movement	preposition	deictic	compound numerative (ordinative quantitative)		
92	is first introduced	verbal	ordinative	verbal		
93	is playing the	verbal group		deictic		
94	it is the	deictic	Verbal	deictic		
101	the sound of	deictic	nominal	preposition		
106	the violin in	deictic	thing	preposition		

Table 5.4.1: NoCo Word Functions

RANK	Lexical Bundle	Function				
		S1	S2	S3	S4	S5
108	use of the	nominal	preposi-tion	deictic		
109	A in the	thing	preposi-tion	deictic		
113	and so on	conjunc-tion	deictic	preposi-tion		
114	and then the	conjunc-tion	adverb-ial	Deictic		
116	as the first	preposi-tion	deictic	ordina-tive		
117	at the beginning of the	preposi-tion	deictic	ordina-tive	preposi-tion	deictic
127	in the introduction	preposi-tion	deictic	thing		
132	is one of	verbal	ordina-tive	preposi-tion		
135	it is a	deictic	verbal	deictic		
139	of the original	preposi-tion	deictic	Epithet/ Nominal		
140	of the second	preposi-tion	deictic	ordina-tive		
143	of this movement	preposi-tion	deictic	thing		
147	similar to the	epithet	preposi-tion	deictic		
153	the idea of	deictic	nominal	preposi-tion		
155	the lower strings	deictic	class-ifier	thing		
161	there is also	deictic	verbal	adverb-ial		

Table 5.4.1: NoCo Word Functions

R A N K	Lexical Bundle	Function				
		S1	S2	S3	S4	S5
164	to develop the	verbal n-fin		deictic		

Table 5.4.2: NoCo LB Word Functions by Slot

Function	S1	S2	S3	S4	S5
adverbial		1	1		
classifier		1			
conjunction				1	
complex conjunction		1			
complex preposition		1			
compound numerative			2		
compound numerative		4			
conjunction	4				
deictic	21	25	13	1	1
post-diectic		1	1		
epithet	1				
epithet/ nominal			1		
nominal	2	4	1	1	
ordinative	1	5	13		
preposition	23	9	12	3	
quantitative	1	1	1		

Table 5.4.2: NoCo LB Word Functions by Slot

Function	S1	S2	S3	S4	S5
thing	1	4	6	1	
verbal	5	4	3		
verbal group	2				
verbal group		1			
verbal n-finite	1				

The number of specific nominals, of Things, is also of interest. Curiously, four of these reference 'violin(s)' and one (with its classifier) references the 'lower strings'. This focus on lower strings and the violin is likely due to their respective roles as the primary bass (i.e., harmonic) and melodic instruments of the orchestra. This suggests that novices are expected to study orchestral music, which again intimates a focus on the Western canon.

Aside from nominals and numeratives, one other observation about word function bears mention here. Unlike the ExCo, the NoCo is largely devoid of word complexes. While there is a complex conjunction, a complex preposition, and a verbal group (all of which are shared with the ExCo), this is the extent of word complexing among these bundles, suggesting either a lack of proficiency with complex lexicogrammatical units or perhaps an insufficient word limit in which to develop ideas more fully.

5.5 Lexical Bundle Disciplinary Content

A characteristic shared with the ExCo is that multiple NoCo LB contain nominals, many of which are specific Things; that is, disciplinary terminology. Table 5.5.1 displays the 21 bundles containing such content (Rank numbers in blue indicate bundles

Table 5.5.1: NoCo LB containing Disciplinary Terms

Rank	Lexical Bundle			Rank	Lexical Bundle			
	S1	S2	S3		S1	S2	S3	S4
1	played	by	the	101	the	sound	of	
14	the	first	move-ment	106	the	violin	in	
25	the	second	move-ment	109	A	in	the	
26	the	third	move-ment	127	in	the	intro-duc-tion	
28	is	played	by	143	of	this	move-ment	
40	of	the	music	155	the	lower	strings	
41	of	the	piece					
46	by	the	violin					
60	the	second	theme					
62	the	theme	and	74	played	by	the	violin
63	the	violin	and	89	in	the	se-cond	move-ment
93	is	playing	the	90	in	the	third	move-ment

shared with the ExCo). Though numerative bundles are counted as an integral part of the Discourse, they are not counted among these nominals since they do not constitute disciplinary terminology, with one caveat: seven of these nominal bundles also contain numeratives as pre-modifiers to the Head of a nominal group. Among the 21 bundles, then, there are 3 4-LB and 18 3-LB. Of these, 5 contain the term 'movement', showing again a concentration on structure. This emphasis is visualized in Word Cloud 5.5.1, wherein the frequency of disciplinary terms is represented by proportional font sizes, thus highlighting the centrality of the term 'movement'.

Word Cloud 5.5.1: Disciplinary Terms by Frequency

https://worditout.com/word-cloud/5510755

In addition to terms related to structure, such as 'movement' and 'theme', several bundles contain disciplinary terms related to performance. This division between structure and performance is mapped in Chart 5.5.1. From this chart the primary concerns of these novice writers is strikingly apparent; they write about structure and support their arguments with examples of what is played or what is indicated in the score. This perspective is further elucidated in Chart 5.5.2, which presents the three areas of music represented by these disciplinary bundles — structure, performance, and phenomenon — as a percentage of the total number of Tokens (171) for all disciplinary-specific bundles. Structure is the most prominent, yet performance is a close second. Phenomenon does not figure prominently mainly because that category is dependent on a single term. It is also noteworthy that the NoCo LB, unlike their ExCo counterparts, are devoid of LB containing centuries, geographic locations, and personal pronouns, suggesting a lack of competence and therefore practice with these disciplinary concerns.

While the relatively limited scope of the NoCo LB may be partially due to the size of the Corpus, it may also suggests that novices are often assigned writing tasks based on the exercise of reading a score while listening and then supporting their arguments about what they hear with examples from the score. Though the NoCo is

limited in size, it is telling that these disciplinary terms are found in both the secondary and tertiary sub-corpora, indicating that similar types of writing assignments persist at both levels. This again indicates that the types of writing expected of novices are not fully aligned with those of experts, who tend to be far more score- than performance-oriented in their writings. This is also supported by the fact that only four of these NoCo disciplinary-specific bundles are shared in common with the ExCo, two of which directly reference structure (i.e., 'movement'), while the other two reference the phenomenon 'music' and 'sound'. This last is the only shared bundle to potentially reference performance.

Chart 5.5.1: Relationship of NoCo Disciplinary Terms
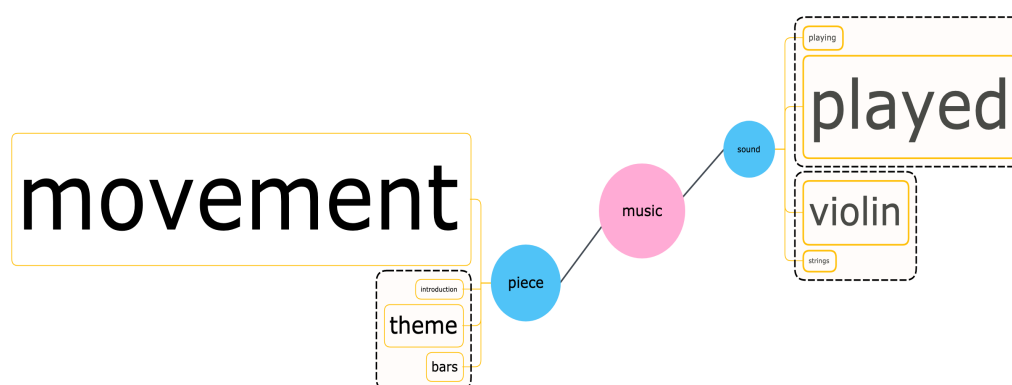
In addition to the various disciplinary-specific LB, there are several bundles containing numeratives. Graph 5.5.1 presents a comparison of the various types of numerative bundles by percentage of their totals in both Corpora, respectively. The novice texts demonstrate a

Chart 5.5.2: Music Areas in Disciplinary LB

5%

51%

44%

● Structure   ● Performance   ● Phenomenon
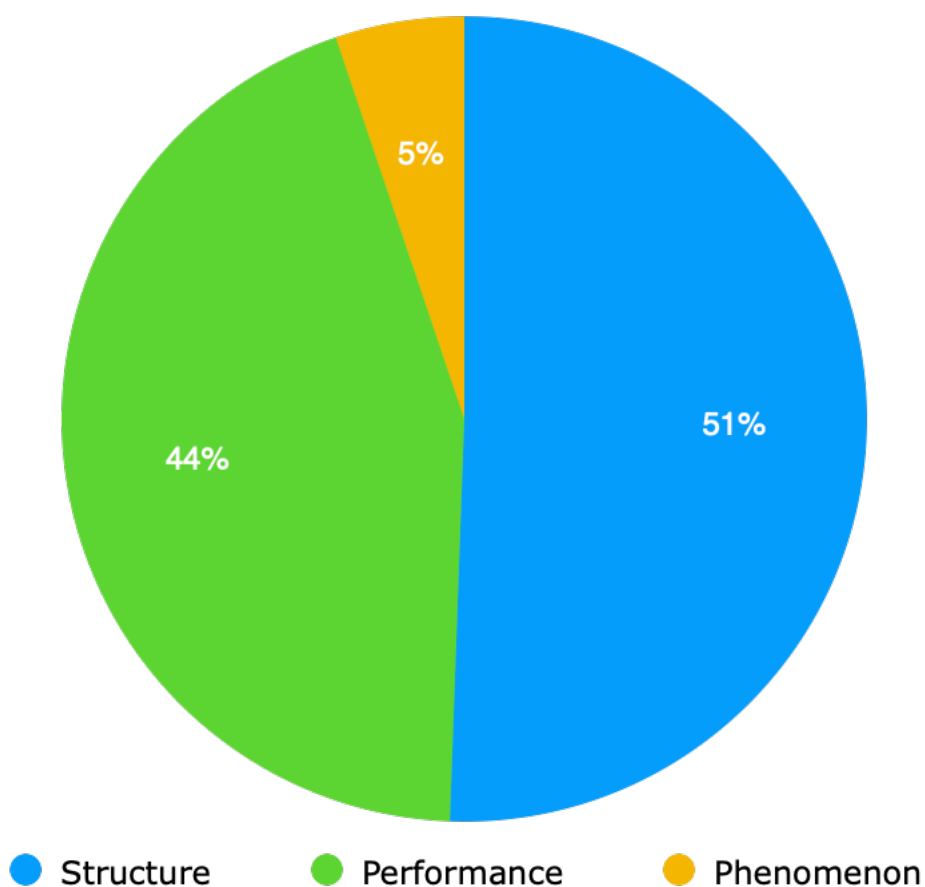
heavier reliance on ordinatives. This is likely due to a focus on structure as it unfolds in time, either as analysis of performances or recordings, of scores, or a combination of both. The far lower percentage of quantitative bundles in the NoCo supports the claim that the novice Discourse is more performance oriented since this type of numerative is more commonly used to discuss portions of, or locations in, a score (e.g., end, beginning), rather than events in a performance. This is further supported by the larger percentage of NoCo compound numeratives as they primarily reference movements, rather than portions thereof.

Graph 5.5.1: ExCo and NoCo Numerative Functions
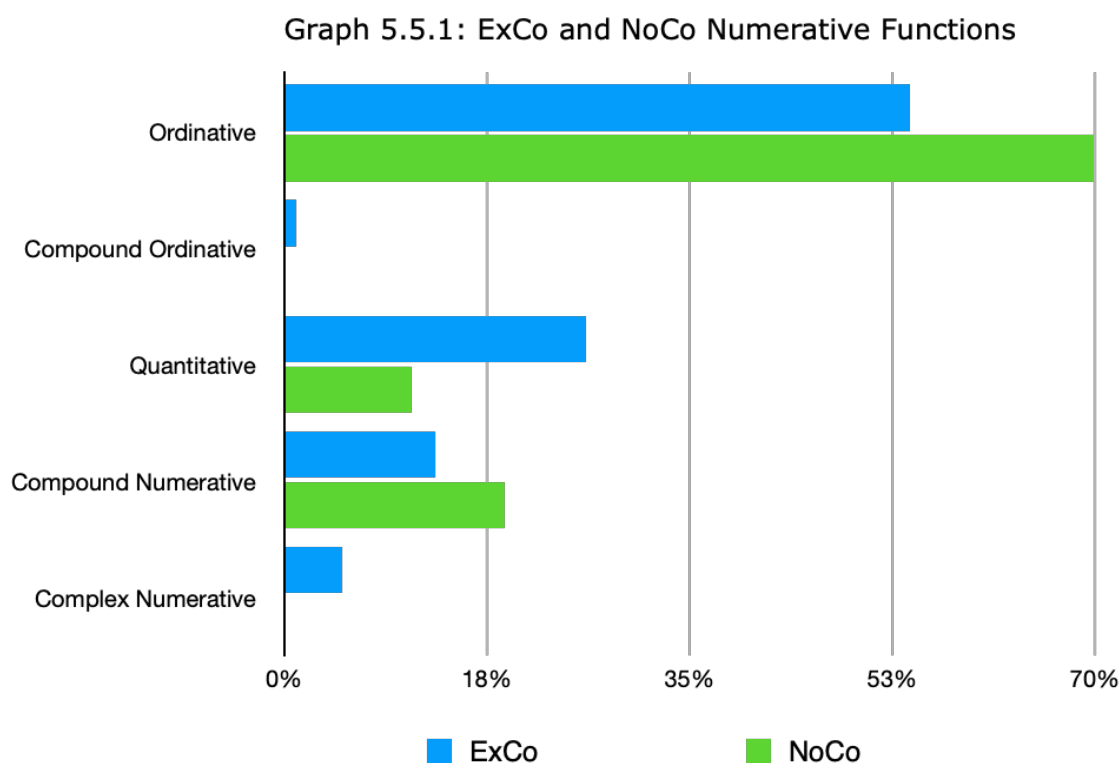
5.6 Prepositions and (Potential) Genitive LB

Having analyzed the word classes and functions for all Novice Corpus LB, it is now possible to examine the distribution of prepositions across the NoCo and isolate those bundles that either form genitives or potential genitives, which will further facilitate discussion of nominals and extended numeratives in the Corpus. Table 5.6.1 lists all the NoCo prepositions by percentage of the total number of prepositions. Following that, Chart 5.6.1 visualizes those percentages. The four most frequent ones account for 87% of all

Table 5.6.1: NoCo Prepositions

Preposi- tion	Fre- quency	% of 46 Total	Preposi- tion	Fre- quency	% of 46 Total
as	1	2%	from	1	2%
as well as	1	2%	in	8	17%
at	5	11%	of	21	46%
between	1	2%	on	1	2%
by	5	11%	with	2	4%

Chart 5.6.1: Proportion of NoCo Prepositions
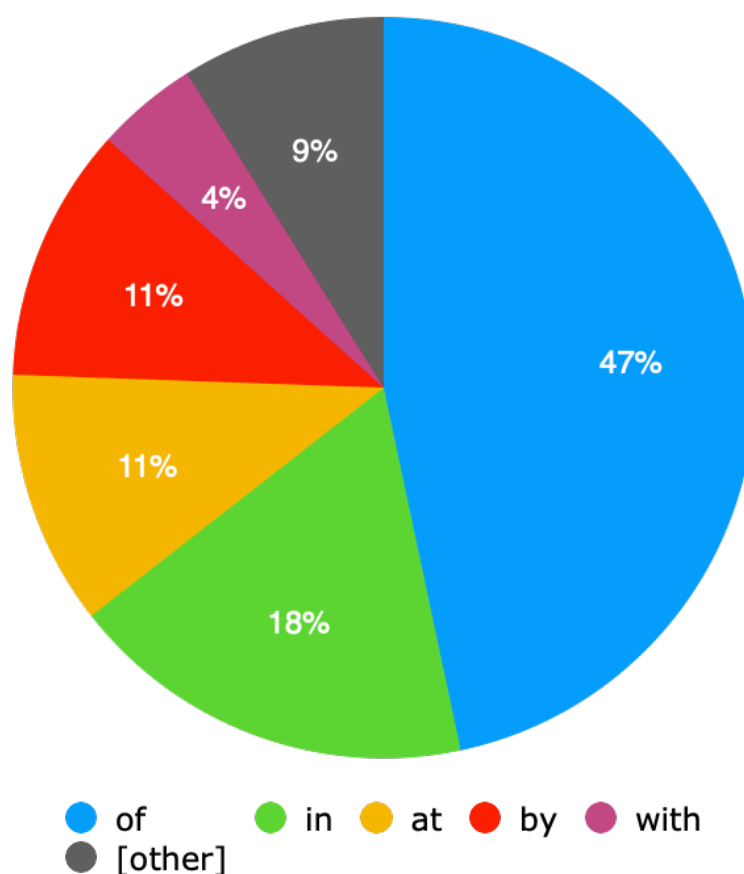
prepositions in NoCo LB: of, in, at, by. The first two of these closely

approximate the percentages found in the ExCo, as shown in Graph

5.6.1. After that, approximation for subsequent prepositions declines, as shown by the third and fourth prepositions, 'at' and 'by'.

Graph 5.6.1: ExCo and NoCo Top Four Prepositions
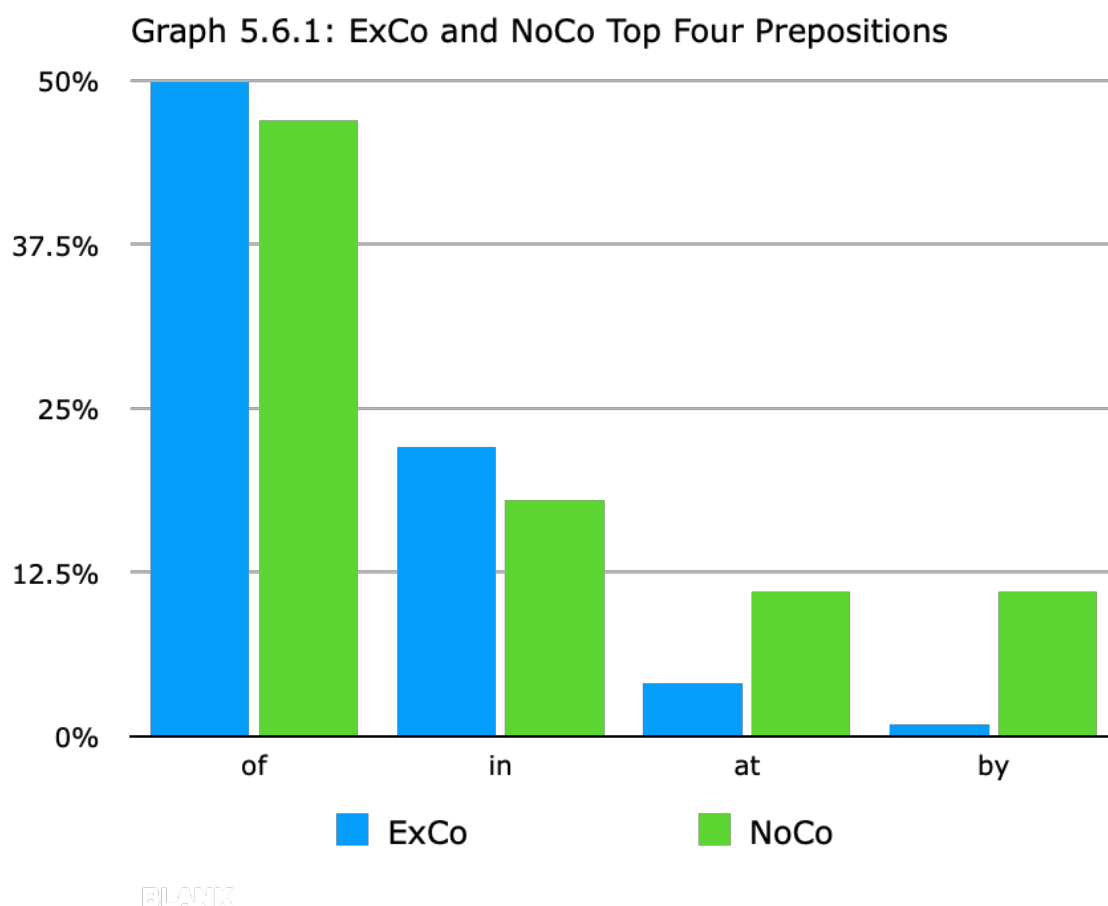
The discrepancy between the frequency of 'by' in both Corpora is generated by its recurrent combination with 'the violin(s)' in the NoCo texts. This is noteworthy for three reasons. Firstly, these bundles often act as metonyms, a single instrument referencing a whole section (i.e., the violins) or a part in the score. Secondly, these bundles show an awareness on the part of the novice writers that performers are generally not named in the Discourse (as in 'the

violinists play…') but rather their instruments are imbued with anthropomorphic agency (as in 'by the violin'). Thirdly, there is no corresponding usage in the ExCo for: 'by' + deictic + [instrument]. Rather, the preposition 'in' is used by experts as a specific reference to an instrumental part in the score, as in ExCo 350 'in the piano'. This form makes it more explicit that a part in the score is being referenced, while the use of 'by' suggests agency or an action that seems to prioritize performance over notation. As noted in the previous chapter, the Expert Corpus prioritizes score analysis well above performance. This discrepancy may then indicate a lack of proficiency with the expectations of the Expert Discourse, unclear expectations from expert instructors, a tendency to employ writing assignments that appear to, or do, prioritize discussion of performance and recordings above score study, or some combination of several or all of these. Whichever combination of the above constitutes the explanation, one seemingly inescapable conclusion is that novices are expected to write in a way that diverges from the norms of the Discourse, which in turn presents them with a barrier to entering the Discourse Community. It is understandable, however, that such granular details regarding prepositional phrase construction are not explicitly brought to the attention of novices, if for no other reason than that their teachers are unaware that bundles form text patterns (Hunston & Francis

301 of 460

2000) that serve in part to construct disciplinary knowledge (Matthiessen 2015). Such a disparity highlights the gate-keeping function that the Discourse performs, as enacted by experts (Gee 2014).

Turning to the most frequent preposition to appear in NoCo lexical bundles, Tables 5.6.2-5.6.4 display all of the bundles that contain 'of' by initial, medial, and terminal slots. There are six LB with initial 'of' (Table 5.6.2), half of which contain disciplinary content and another third of which contain numeratives. The group of five bundles with medial 'of' (Table 5.6.3) includes the only 5-LB and one of the 4-LB. Three of these bundles are portions and two are quanta. Of the bundles with terminal 'of' (Table 5.6.4), there are nine dissociated Heads of extended numeratives, half of which are portions, and then one each of aggregate, facet, make-up, and quantum.

Table 5.6.2: NoCo LB Initial 'of'

Rank	Slot 1	Slot 2	Slot 3	Type
139	of	the	original	[other]
40	of	the	music	disciplinary
41	of	the	piece	disciplinary
143	of	this	movement	disciplinary
39	of	the	first	numerative
140	of	the	second	numerative

302 of 460

Table 5.6.3: NoCo LB Medial 'of'

Rank	S1	S2	S3	S4	S5	Type
13	one	of	the			quantum
18	end	of	the			portion
22	the	end	of	the		portion
108	use	of	the			[other]
117	at	the	beginning	of	the	portion

Table 5.6.4: LB Terminal 'of'

Rank	S1	S2	S3	S4	Type
4	the	end	of		portion
15	the	use	of		[other]
16	a	lot	of		portion
29	the	beginning	of		portion
34	at	the	end	of	portion
44	at	the	beginning	of	portion
64	a	group	of		aggregate
101	the	sound	of		facet
132	is	one	of		quantum
153	the	idea	of		make-up

In total, there are 21 LB containing the preposition 'of'. Of these, 15 function as numeratives, 13 of which form extended numeratives. The latter divide into various types, all of which are visualized in Chart 5.6.2 as percentages of the total 15 numerative bundles. Portion bundles account for half of these, again suggesting a focus

on large sections of a piece, perhaps through hearing a performance or reading a score.

Chart 5.6.2: NoCo Numeratives (% of 15 LB)

Numerative 14% **Aggregate** 7% **Portion** 50% **Quantum** 14% **Facet** 7% **Make-up** 7%
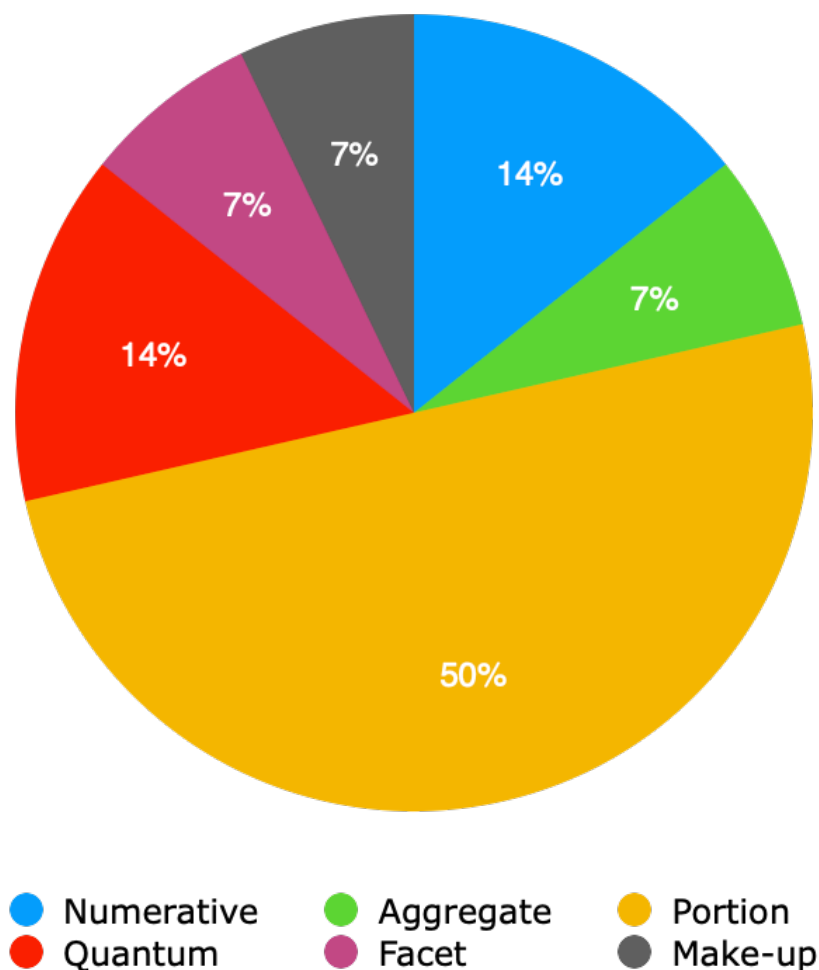
BLANK

5.7 Lexical Bundle Discourse Functions

The results of the analysis employing Biber et al.'s Taxonomy of LB discourse function (2004) are listed in Table 5.7.1 along with the percentage of the total 65 LB Types of each of the four discourse Types employed herein (i.e., Discourse, Propositional, Referential, Stance) as shown in Chart 5.7.1.

Referential bundles account for more than half (58%), indicating the potential for ambient slots to be filled with propositional content. The next largest functional group is the Propositional LB (34%) containing terminology. The remaining two groups, Discourse and Stance, account for a mere 8% between them, with Stance being represented by a single bundle, NoCo 55 'can be seen'. This finding suggests that novices are either unfamiliar with, or unpracticed at, hedging and boosting strategies manifested through Stance bundles. These findings contradict Nekrasova's (2009) claim that referential LB are often neglected by novices, suggesting instead that, at least, these novices lack facility with discourse and stance, in particular. This likely points to a specific lacuna in writing instruction, one that again may not be well understood by disciplinary instructors.

A comparison of these NoCo percentages with the ExCo ones is presented in Graph 5.7.1. This comparison reveals a substantial discrepancy between the use of Propositional and Referential LB, with novices relying far more heavily on the former and rather less on the latter. There are several possible explanations for this difference. Firstly, shorter texts are less likely to contain numerous referential bundles as space is considerably more limited than in a

Table 5.7.1: NoCo LB Discourse Functions

Rank	Frequenc	Range	Bundle	D-Type
57	7	4	in order to	Discourse
70	6	4	contrast with the	Discourse
84	5	4	as well as	Discourse
113	4	3	and so on	Discourse
1	27	7	played by the	Propositional
14	14	4	the first movement	Propositional
25	11	6	the second movement	Propositional
26	11	5	the third movement	Propositional
28	10	6	is played by	Propositional
40	9	7	of the music	Propositional
41	9	5	of the piece	Propositional
46	8	4	by the violin	Propositional
49	8	5	in bars [#] and	Propositional
60	7	4	the second theme	Propositional
62	7	4	the theme and	Propositional
63	7	4	the violin and	Propositional
74	6	3	played by the violin	Propositional
89	5	4	in the second	Propositional
90	5	3	in the third movement	Propositional
93	5	4	is playing the	Propositional
101	5	3	the sound of	Propositional
106	5	3	the violin in	Propositional
109	4	3	A in the	Propositional
127	4	3	in the introduction	Propositional
143	4	3	of this movement	Propositional
155	4	3	the lower strings	Propositional
4	19	12	the end of	Referential

Table 5.7.1: NoCo LB Discourse Functions

Rank	Frequenc	Range	Bundle	D-Type
8	16	8	in the first	Referential
10	16	8	the same time	Referential
12	14	8	at the same time	Referential
13	14	8	one of the	Referential
15	14	10	the use of	Referential
16	13	7	a lot of	Referential
17	13	11	at the end	Referential
18	13	7	end of the	Referential
22	12	7	the end of the	Referential
27	11	3	there is a	Referential
29	10	6	the beginning of	Referential
32	9	4	and it is	Referential
34	9	9	at the end of	Referential
37	9	6	in the second	Referential
38	9	5	in the third	Referential
39	9	7	of the first	Referential
44	8	5	at the beginning of	Referential
64	6	4	a group of	Referential
69	6	4	between the two	Referential
72	6	5	from the first	Referential
82	6	5	with the first	Referential
83	5	3	and this is	Referential
86	5	3	followed by a	Referential
92	5	4	is first introduced	Referential
94	5	4	it is the	Referential
108	5	5	use of the	Referential
114	4	4	and then the	Referential

Table 5.7.1: NoCo LB Discourse Functions

Rank	Frequenc	Range	Bundle	D-Type
116	4	4	as the first	Referential
117	4	3	at the beginning of the	Referential
132	4	4	is one of	Referential
135	4	4	it is a	Referential
139	4	4	of the original	Referential
140	4	4	of the second	Referential
147	4	3	similar to the	Referential
153	4	3	the idea of	Referential
161	4	4	there is also	Referential
164	4	3	to develop the	Referential
55	7	4	can be seen	Stance

full-length journal article. Secondly, novice use of propositional bundles is likely a learned response to the prompt for any short writing assignment and may even be fueled by instructor's feedback insisting on greater clarity of expression, as such comments may encourage greater reliance on terminology. Thirdly, this may also be an indication of a lesser degree of lexicogrammatical complexity in novice writing. Lastly, there is the tantalizing possibility that this discrepancy can be explained by the greater focus on music-as-phenomenon by novices. As seen in the previous chapter, experts focus rather more on scores, while it was noted above that novices appear to focus on performance. If that is the case, it could mean

Graph 5.7.1: ExCo and NoCo LB Discourse Functions
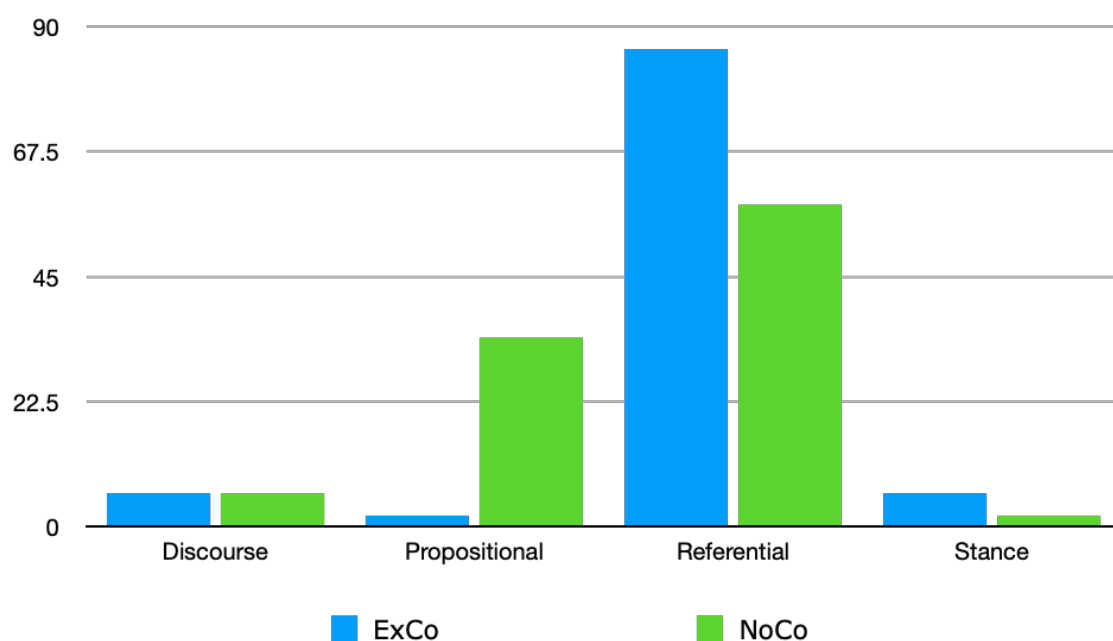
Chart 5.7.1: NoCo LB Discourse Functions
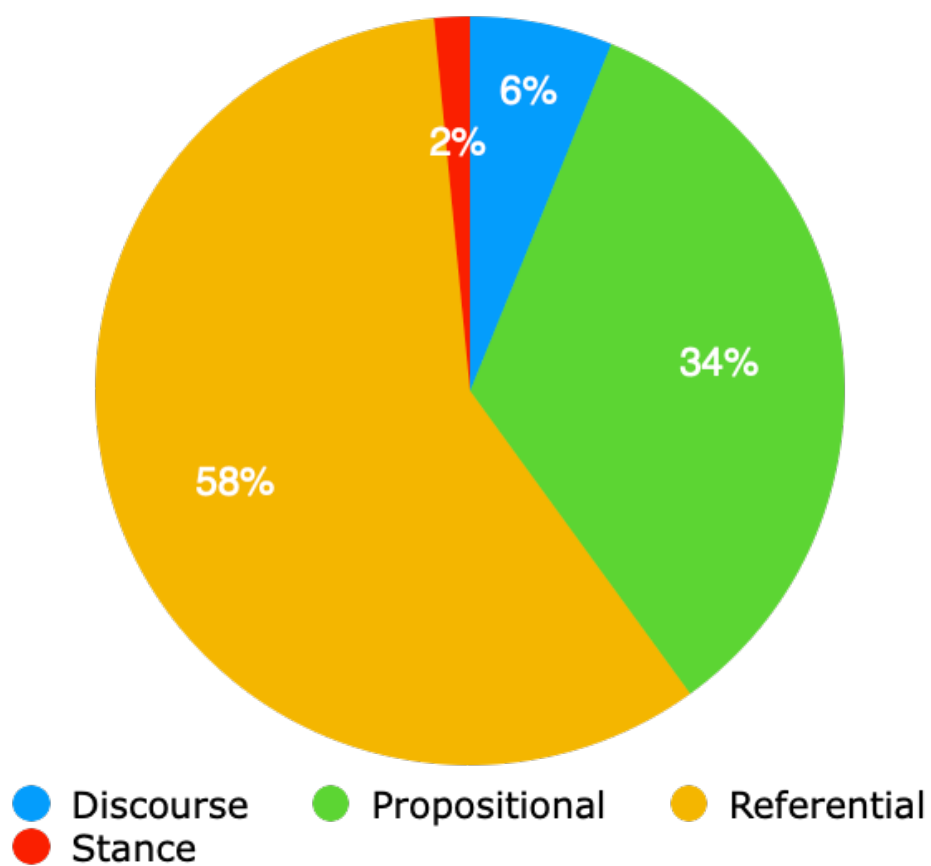
that novice writing tends to be more focused on music than expert writing, with its emphasis on extramusical context.

5.8 Global Views of Complete Texts

In preparation for examining the global views of complete NoCo texts, the Global View Color Coding Key 5.8.1 is presented again below. Global View 5.8.1 presents four complete texts, two each from the Secondary and Tertiary sub-corpora, one per panel of the View, beginning with the Secondary texts (upper row) and continuing with the Tertiary (lower row).

Key 5.8.1: Global View Color Coding

Text

INTRODUCTION

BODY

[QUOTATION]

SUMMARY/CONCLUSION

NON-DISCIPLINARY LB

DISCIPLINARY LB

There are numerous points of interest in these four global views. Internally, the Secondary and Tertiary texts display markedly different patterns. The most obvious of these is the difference in

Global View 5.8.1: Four Novice Essays
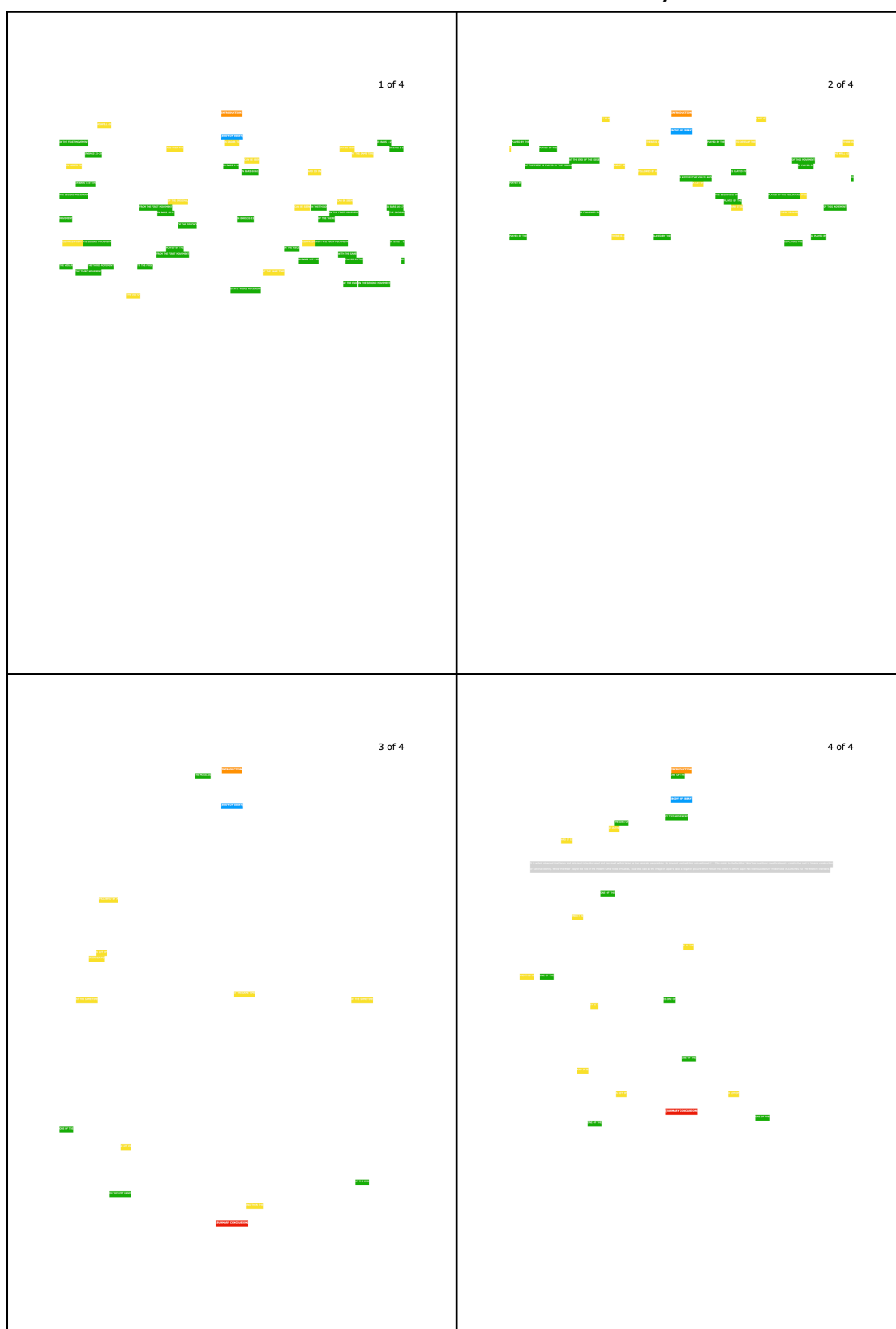
text length, the former being less than half to a third the length of

the latter. This is to be expected as higher levels of education tend

to assign longer and more advanced writing tasks. Far more interesting, however, is the substantially different use of lexical bundles. The Secondary texts are replete with disciplinary-specific bundles (highlighted in green), yet the same are largely absent from the Tertiary texts. There is also a higher density of non-disciplinary bundles (highlighted in yellow) in the Secondary texts (though still rather less than the disciplinary LB). In fact, there is a scarcity of bundles altogether in these two Tertiary texts. At first, this might seem to suggest a problem in the construction of the Corpus, such as an imbalance between the two sub-corpora. Indeed, there is an imbalance but it favors the Tertiary sub-corpora, which contains 20,806 words compared with 7,642 in the Secondary one (Tables 3.2.1-3.2.2). In other words, the greater number of words in the Tertiary sub-corpus should account for many of the LB found in common with the Secondary texts, particularly given that '[i]n academic prose, about 21% of the words occur in a recurrent lexical bundle' (Longman 1999:995). Thus, other explanations for the discrepancy in the number of LB per complete text are more probable.

The first possible explanation is that the complete texts themselves show that the Tertiary essays are neither as score-based nor as listening-based as the Secondary ones. For instance, one such text

is about music for a video game and how that music furthers the action of the game. In this instance, the novice is tasked with writing something that is not focused on music-as-phenomenon (as is also largely the case with Expert Discourse), yet also not focused on performance or a score. The result is a text that begins to approach the expectations of expert writing though without its lexicogrammatical apparatus. This observation then leads to a second plausible explanation. These novices were first confronted with the task of writing complete academic texts in English at the tertiary level, whereas their secondary counterparts had begun that process at an earlier stage. As younger language learners, they may have absorbed highly recurrent word strings more readily than even a slightly older learner could.

There are a few other points to note about the Global View of these texts. The first is structural. The introductions of these texts are brief and therefore largely devoid of bundles. As for the Secondary texts, both lack a discrete Summary/Conclusion section, no doubt on account of their brevity. In this same section in the Tertiary texts, there are no bundles in the first and only one in the second. Again, this is likely on account of brevity, but also is somewhat reminiscent of the ExCo Global Views in which the bookending sections contain far fewer LB than the body of the text. Lastly, the

second Tertiary text exhibits one other feature in common with the ExCo texts: the presence of a quotation.

Before leaving this section, it is worth noting two drawbacks to the above Global Views: 1) they only present a highly limited number of texts, so extrapolations based on them must be treated cautiously; and 2) the novice texts may actually contain ExCo LB not frequently recurrent in the NoCo. If that is the case and those were to be taken into account, it is possible that these Global Views would reveal a greater density of bundles per text.

5.9 Cartography of NoCo Ambient Content

The 318 Tokens of the 65 NoCo LB have 636 ambient slots. Of these, 372 are filled with nominal content, constituting 58% of the total potential slots. As with the ExCo, the P-slot is considerably less productive than the S-slot; in fact, it is precisely half so at 124 filled slots of the 248 S-slots. Table 5.9.1 presents these numbers alongside those of the ExCo ambient content. Interestingly, the P-slot is 10% more productive in the NoCo while the S-slot is 14% less so than in the ExCo. Though neither of these represents a large discrepancy, the greater size of the latter discrepancy may be explained as a consequence of the significantly larger proportion of referential bundles in the ExCo. In turn, this suggests a need for

novices to gain fluency with a broader array of referential strategies for the purpose of increasing the saturation of propositional content in their writing. Contrary to earlier findings, this then reharmonizes with Nekrasova's (2009) statement that referential LB are often neglected by novices and demonstrates that, though they may account for 50% of all LB Discourse Functions, Referential bundles are still less utilized in novice writing than in the Expert Discourse. There is another possible explanation, however, for the lower productivity of the NoCo S-slots: they may be partially filled with grammatically awkward constructions that prevent them from being analyzed according to the rubric employed herein.

Table 5.9.1: ExCo and NoCo Ambient Content by Slots

	ExCo		Noco	
	No. of Filled Slots	% of Total 17,837 per side	No. of Filled Slots	% of Total 318 per side
P-slot	5,179	29%	124	39%
S-slot	16,456	92%	248	78%

Table 5.9.2 compares the categories of ExCo and NoCo ambient content by frequency and percentage. The three largest categories for each are highlighted in pink. For the NoCo, those three account for 59% of the total content. In decreasing size they are 'structure', 'instrument', and 'extramusical'. This contrasts notably with the

Table 5.9.2: ExCo and NoCo Ambient Content Categories

Ambient Category	ExCo Frequency	% of 22,711	NoCo Frequency	% of 372
Articulation	30	>1%	3	>1%
Composer	165	>1%	7	2%
Composition	2,586	11%	18	5%
Contrast			5	1%
Counterpoint	149	>1%	15	4%
Dynamic	68	>1%	9	2%
Emotion	29	>1%	3	>1%
Extramusical	10,472	46%	40	11%
Harmony	754	3%	11	3%
Instrument	160	>1%	52	14%
Instrumentation	167	>1%	2	>1%
Media			4	1%
Melody	533	2%	6	2%
Notation	219	>1%	12	3%
Numeral			2	>1%
Ornamentation	34	>1%	2	>1%
Part	633	3%	10	3%
Performance	455	2%	1	>1%
Performer	222	>1%	4	1%
Pitch	738	3%	3	>1%
Recording	151	>1%	1	>1%
Register			1	>1%
Rhythm	765	3%	11	3%
Sound	100	>1%	1	>1%
Structure	1,899	8%	125	34%
Style	200	>1%	3	>1%
Technique	28	>1%	4	1%

316 of 460

Table 5.9.2: ExCo and NoCo Ambient Content Categories

Ambient Category	ExCo Frequency	% of 22,711	NoCo Frequency	% of 372
Tempo	50	>1%	3	>1%
Texture	67	>1%	7	2%
Timbre	30	>1%	1	>1%

ExCo, in which the three largest categories are 'extramusical', 'composition', and 'structure'. A comparison of percentages for these largest categories between the Corpora is shown in Graph 5.9.1. The three largest ExCo categories account for 65% of its total ambient content. At 59% of the total, the top three NoCo categories account for nearly the same percentage as the ExCo; but that is where the similarity ends. As the graph reveals, there are very different concerns at the core of the Expert and Novice texts. While the former is dominated by extramusical concerns, the latter is substantially focused on musical structure, and therefore at least in part on score notation, though also on performance, as indicated by the 'instrument' category, which accounts for less than 1% of ExCo content. While specific compositions do account for 5% of NoCo ambient content, that is less than half their presence in the ExCo. Perhaps most telling, however, is the considerably smaller percentage of the 'structure' category in the ExCo than in the NoCo. As previously observed, a significant proportion of the ExCo LB contain disciplinary terms centered on structure, and analysis of

Graph 5.9.1: ExCo and NoCo Top Ambient Categories
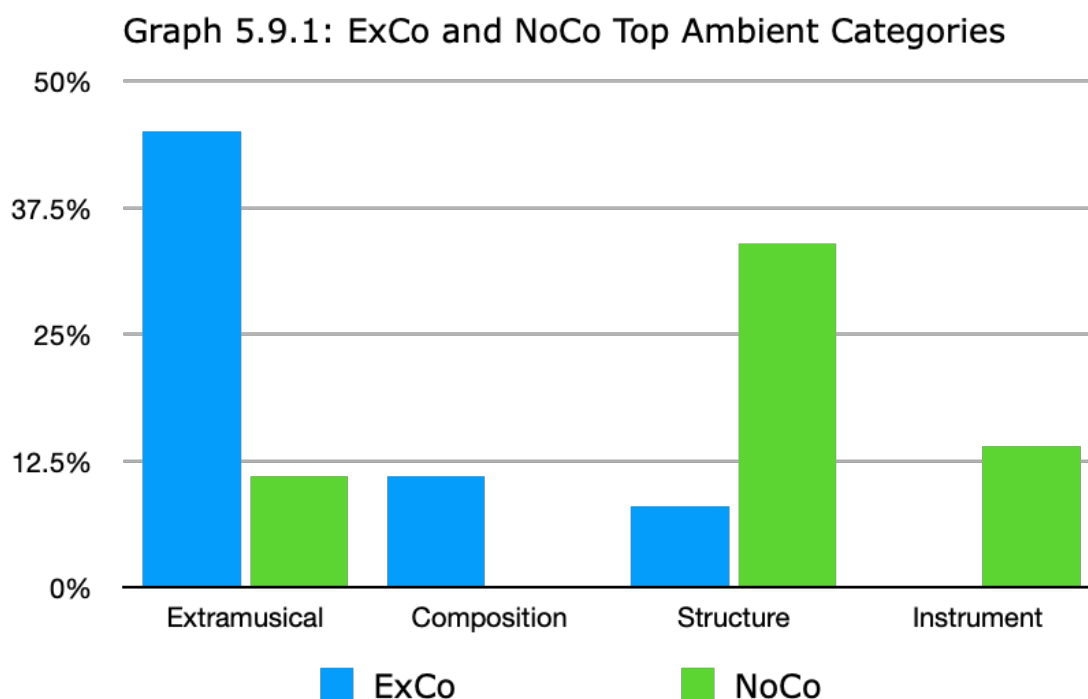
structure accounts for one of the major concerns of expert writers. In light of the percentages for ambient categories, specifically the largest in each Corpus (i.e., ExCo 'extramusical' and NoCo 'structure'), it appears that experts perceive their Discourse as one dedicated to analysis when in fact it is largely given to contextualization, as evidenced by the overwhelming presence of extramusical content, and most especially to historical contextualization, as evidenced by the large number of *century* bundles and *century* content in the ambient slots. Two of the NoCo Tertiary texts reference centuries in the S-slots of six LB Tokens: sixteenth century is mentioned three times in one text, and eighteenth and nineteenth centuries are mentioned once and twice,

respectively. Because there are only two such novice texts, this does not make for a productive comparison with expert writing; however, the lack of additional Tokens in the NoCo suggests that novices tend not to focus on historical contextualization to nearly the same extent as experts. Moreover, the primary emphasis on structure in the NoCo indicates that expert instructors view this aspect of music as central to novice education. This focus then distorts both expert and novice understandings of the Discourse. This claim is further supported by the lack of a '[quotation]' category among the NoCo content, as that category is vital to historical contextualization , which is dependent on primary sources.

Chart 5.9.1 visualizes the NoCo ambient content by percentage and Chart 5.9.2 organizes those categories into their musical relationships. The first chart demonstrates the prominence of the 'structure' category, while the latter maps the other categories that are grouped together with structure (Chart 5.9.2 Magnification 1). There are only a few of these additional categories, among which counterpoint, rhythm and harmony — all fundamental aspects of music — figure most prominently. This area of the map suggests connections both to a score and to performance. Setting aside the extramusical category, the next largest concentration of categories are those grouped together around sound, most notably the

Chart 5.9.1: NoCo Ambient Content Categories
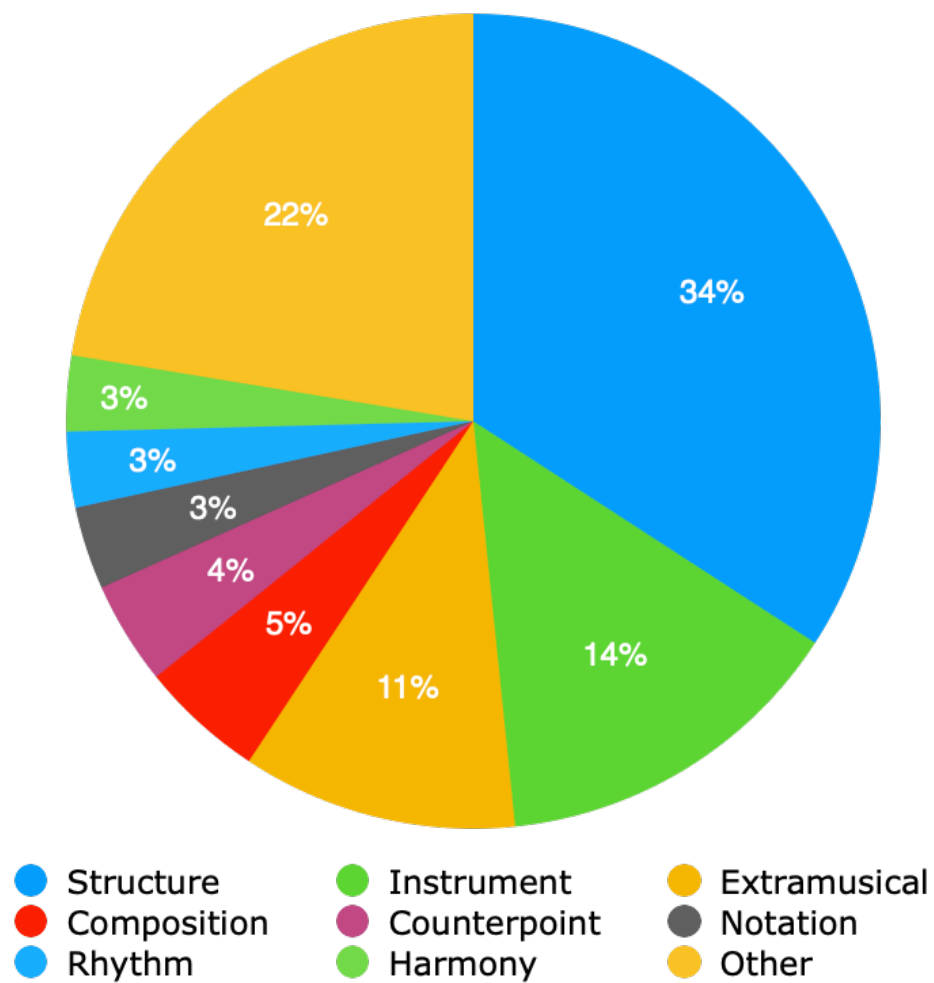
substantial instrument category (Chart 5.9.2 Magnifications 2a and 2b). All the categories gathered here are used to describe performance, further demonstrating the NoCo prioritization of this over structure and score notation.

Map 5.9.1: NoCo Ambient Content Category Relationships
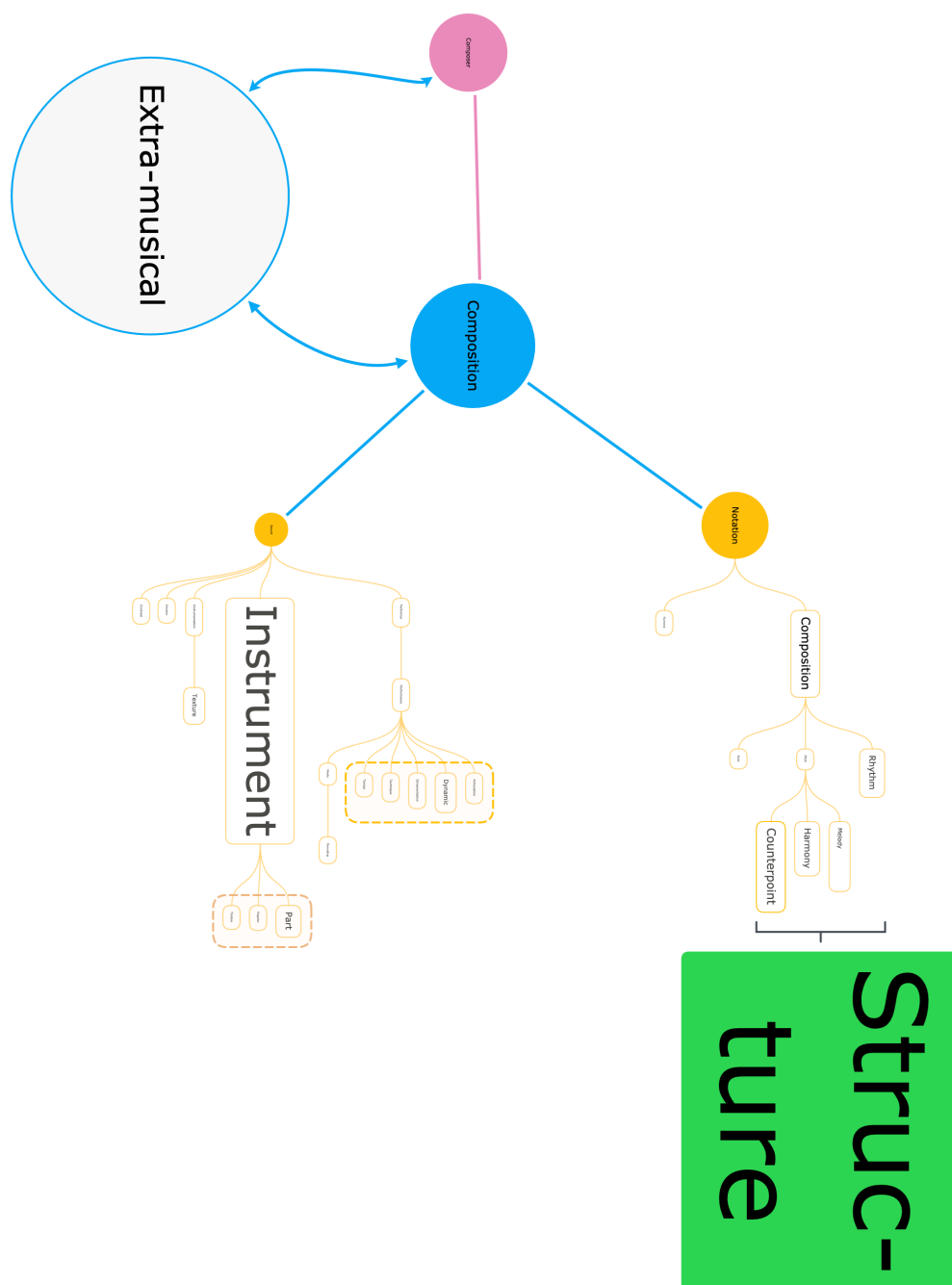
Chart 5.9.2 Magnification 1: Structure

Notation

Numeral

Composition

Style

Pitch

Rhythm

Counterpoint

Harmony

Melody

St
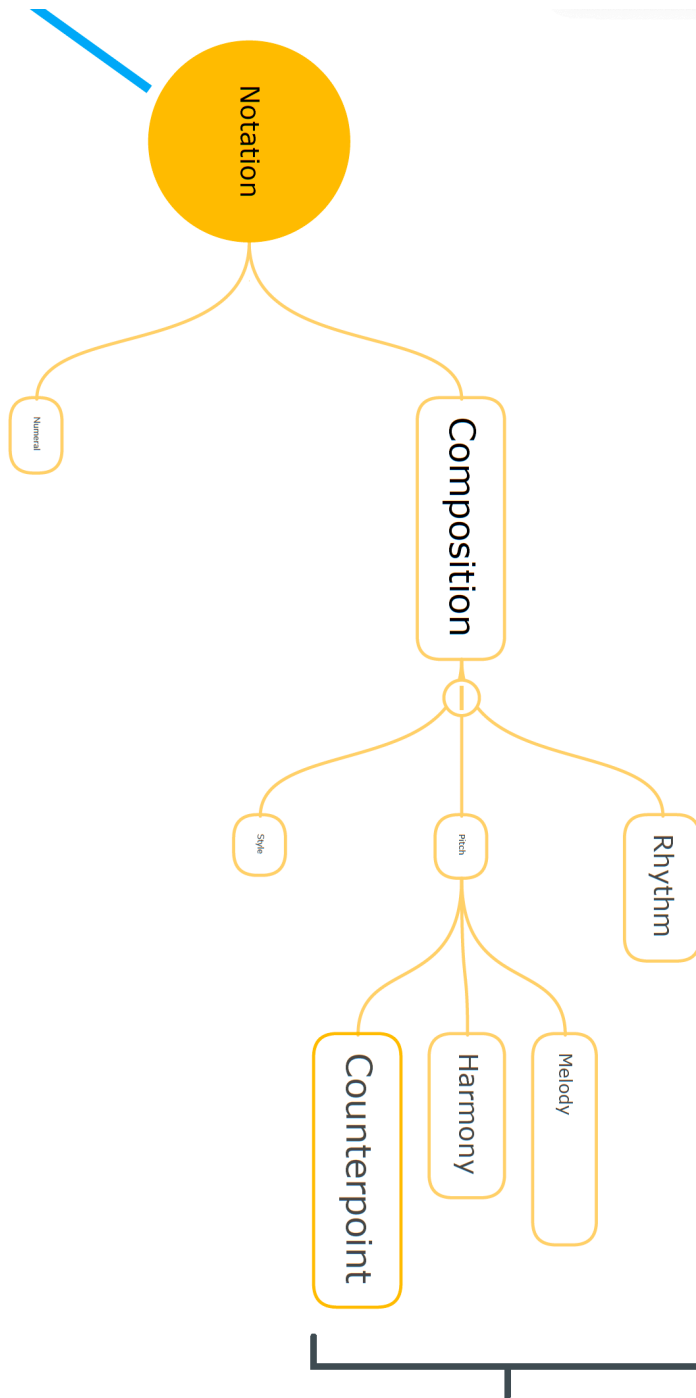
Chart 5.9.2 Magnification 2a: Sound

Sound

Instrument

Performer

Performance

Media

Recording

Tempo

Technique

Ornamentation

Dynamic

Articulation

Timbre

Register

Part
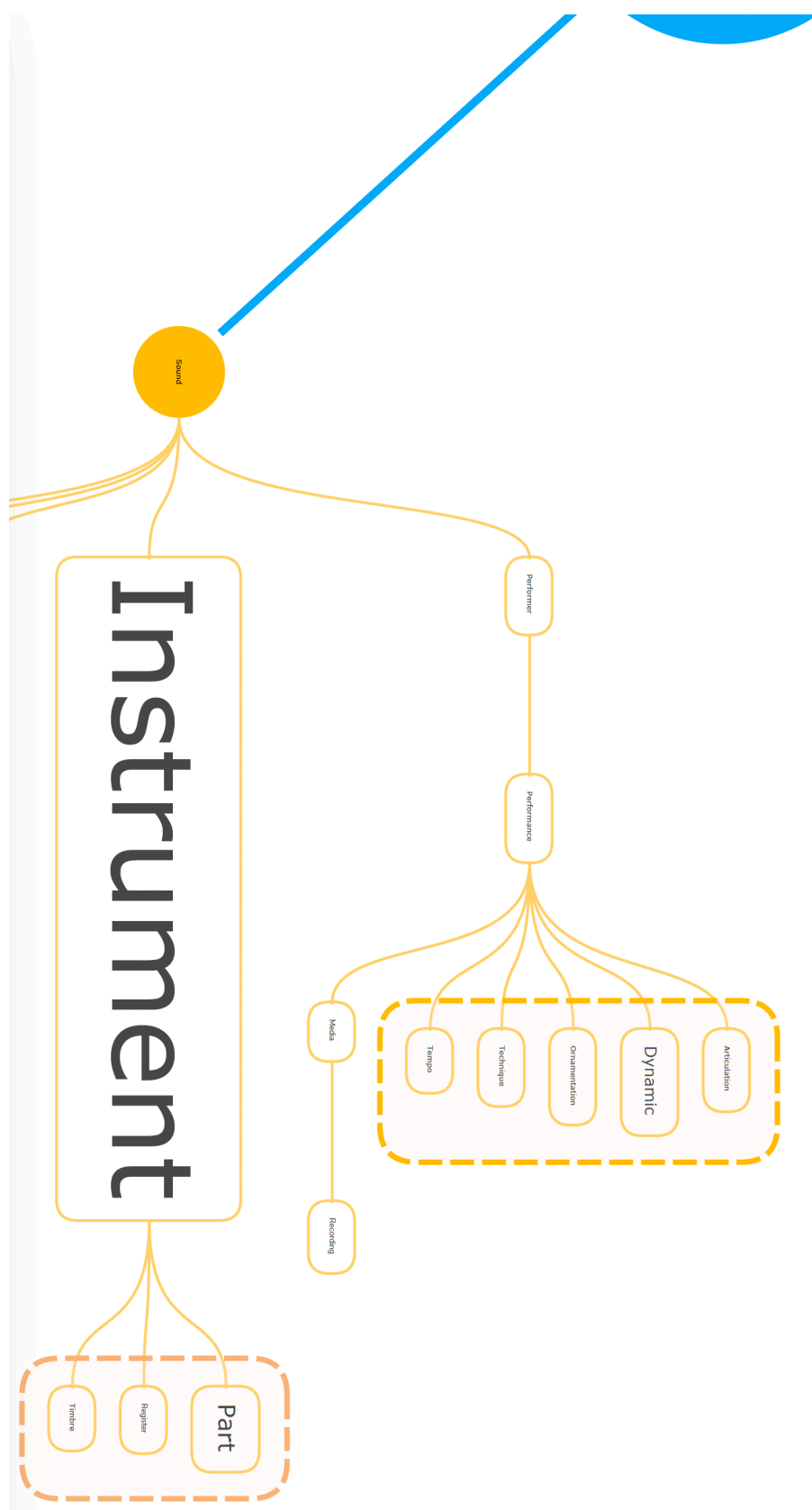
Chart 5.9.2 Magnification 2b: Sound

Sound

Contrast

Emotion

Instrumentation

Texture

Instrument

Timbre

Register

Part
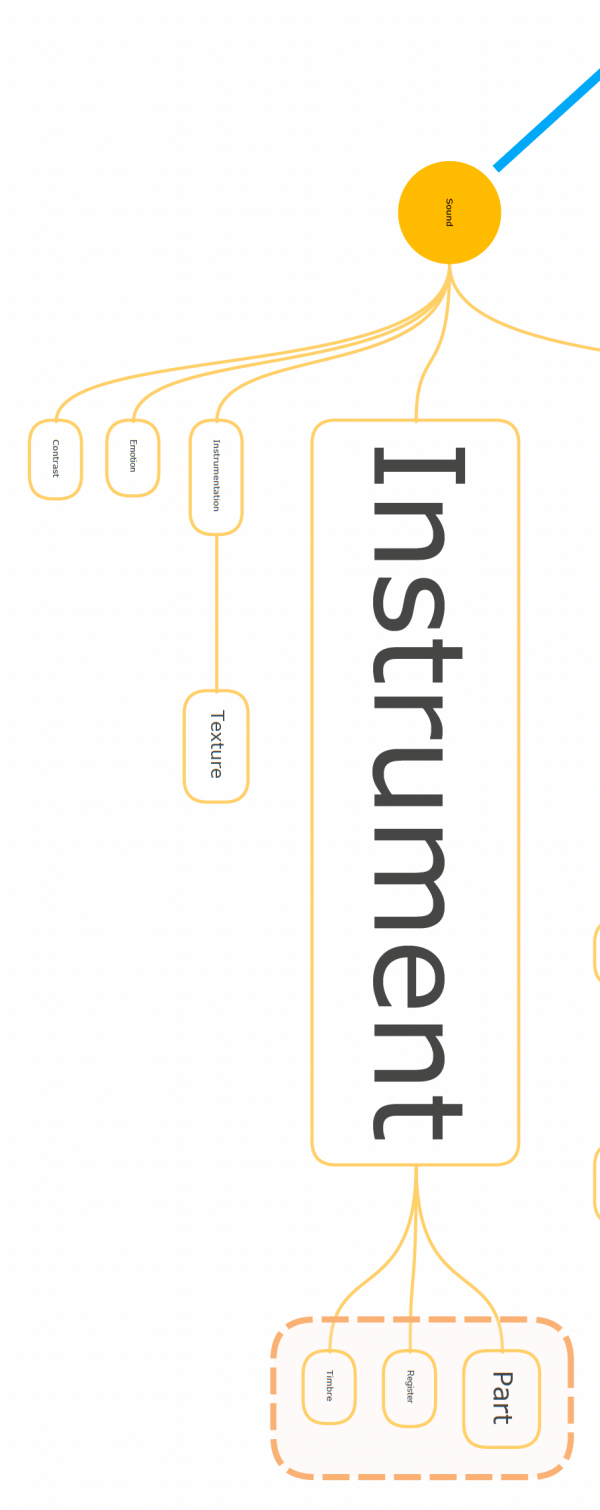
5.10 Summary of Findings

Based on the preceding comparison of findings from both Corpora, it is striking the degree to which novices exhibit acquisition of certain aspects of the Discourse. This is all the more remarkable given that the texts in the NoCo were all produced by L2 learners. Indeed, there seems to be little evidence in the Novice Corpus that overtly identifies those texts as the products of L2 writers, save perhaps for less lexicogrammatical variety than might be expected, particularly from tertiary-level novices, as suggested by the large number of NoCo Syntagmatic LB Types not found in the ExCo.

As to which aspects of the Discourse these novices have acquired, they clearly have found value in the use of highly recurrent bundles, with 60% of the NoCo bundles being shared with the ExCo. This is to be expected not only as an inevitable statistical outcome but also as a learning strategy (Wray 2008), conscious or otherwise. Such is likely the case, for instance, with numerative bundles, which figure conspicuously in both Corpora. However, more granular analysis reveals that different types of numeratives and extended numeratives are employed by the two groups of writers (despite the comparable use of the structural/genitive 'of' in both Corpora), demonstrating a deviation by novices from the norms of Expert Discourse. As examples of this, ordinative and portion bundles

appear considerably more often in the NoCo, demonstrating a discoursal focus on music as phenomenon unfolding in time. Likewise, the noticeably smaller proportion of quantitative bundles reveals a lesser concern with location in a score. Aside from discrepancies in the percentage of particular numeratives employed, the NoCo disciplinary-specific LB demonstrate a concern for performance with terminology such as 'sound', 'played', 'playing', 'strings', and 'violin'. Furthermore, there is a 32% greater use of propositional LB than in the ExCo, where this type is negligible. This appears to be especially true of the Secondary novice texts, as seen in the Global View. While there are many LB containing structural terminology, such as 'movement', these appear to serve discussion of performance more than analysis of a score, given the lack of terminology for specific portions of scores.

Finally, there is much less extramusical ambient content in the NoCo than in the ExCo. This demonstrates a different set of priorities than those evident in the ExCo, which explains why there are no NoCo bundles containing centuries, geographic locations, or personal pronouns; the NoCo texts are not concerned with socio-historical contextualization of music, but rather with describing performance through musical structure. In this sense, novice writers fall far short of the gate-keeping expectations set for them by the Discourse Community. That said, it is apparent that their instructors assign

writing tasks that do not precisely align with the priorities of the Expert Discourse, thus making it impossible for novices to acquire a fuller understanding of the Discourse, its constituents, and its purpose.

6 Conclusion

6.1 Introduction

Employing a corpus-based methodology focused largely on frequency, the present study set out to map Music Discourse, narrowly defined herein as writings of the Academic register produced by expert and L2 novice writers. This required the construction of two bespoke Corpora, one to represent each of these two groups of writers. Corpus is particularly useful for such comparative research because it employs authentic materials (Granger 2009) for study under specific conditions that permit inferences of a larger reality (Stubbs 2007a), the results of which are replicable and verifiable (McEnery et al. 2006), and beneficial to L2 novices (Gilquin 2015; Granger et al. 2015).

The goal of the present comparison of expert and L2 writing has been to determine the degree to which novices have attained competence in the Discourse as practiced by experts. To ensure a balanced and representative comparison, it was necessary to select appropriate texts with great care. Thus, journal articles were chosen for the Expert Corpus since they were readily accessible, offered the opportunity to create an appropriately sized corpus from a large selection of texts, and could be selected from the most widely cited, and therefore most prestigious, publications within the Music

Discourse Community. In other words, it matters who writes these articles and who reads them. These articles are produced by scholars working in the related disciplines of Ethnomusicology, Musicology, and Music Theory, all of which may generally be grouped together under Musicology. While many novice writers intend to pursue further study and careers as performers rather than scholars, they are still required to read selections of this Discourse as part of their studies. Moreover, they are assigned writing tasks intended to teach them how to write in this manner, tasks that often take the form of an essay. Thus, it is ultimately hoped that the findings from this study will aid not only novices in the process of acquiring discourse competence, but also experts tasked with the work of teaching the mechanics of the Discourse to them.

6.2 Summary of Research

Since analysis must seek to describe interweaving patterns that form a complete discourse, rather than merely label its separate parts (Sinclair 2004b), the methodology of this study was designed to investigate not only formulaic items (i.e., lexical bundles), but also the original content with which these bundles intersect. In this way, the configurations of propositional content that constitute Music Discourse could be mapped. This necessitated a focus on

nominal content, as nominals are the primary word function by which experiential meaning is construed (Thompson 2014:40). Other aspects of the Discourse, such as lower frequency instances of originality, were excluded from the analysis as they could not produce sufficient quantities of data from which to generalize. By focusing on disciplinary content, as instantiated both through the repetition of certain lexical bundles and their connections to ambient content, it was possible to generate a cartography of the Discourse that provides an overview of its major concerns. This was accomplished by generating a taxonomy of ambient content based on manual analysis of every preceding and succeeding ambient slot for each Token of all high-frequency lexical bundles. The repetition of configurations of bundles with certain categories of ambient content revealed the features typical of, and patterns unique to, Music Discourse. One proof of the reliability of this methodology lies in the fact that the results were regrettably predictable. That is, the Discourse remains grounded in the same concerns and investigative procedures it has pursued since its inception in the nineteenth century, the Discourse community's conviction of the contrary notwithstanding. More revealing is the demonstration that expert instructors assign novices writing tasks that are not fully aligned with the practices of the Expert Discourse. Though both expert and novice writing exists within a single register, the Academic Register,

they serve different purposes; thus, they constitute different yet related genres, as genre is:

> …register plus purpose. That is, it includes the general idea of what the interactants are doing through language, and how they organize the language event, typically in recognizable stages, in order to achieve that purpose. (Thompson 2014:42-43)

Therefore, a basic understanding of the generic outlines of these Discourses is useful for understanding what binds their various texts into a single genre.

Since the majority of the ExCo LB either cross functional group boundaries or bisect a group, it was necessary to analyze them both by word class (i.e., syntagmatically) and word function to obtain a more complete understanding of their role in the lexicogrammar and the Discourse. Naturally, the latter determination could only be made by examining the ambient content connected to each bundle Token. In this manner, it was possible to study the construal of propositional content in the Discourse to discover its most frequent patterns. Thus, this global perspective demonstrates the

epistemological concerns of the Discipline, over and against the claims of its practitioners.

The Expert Discourse exhibits a few features that distinguish it as a particular sub-register of academic writing. The most obvious of these are the presence of disciplinary lexical bundles. These are most prominent among the numerous 3-LB, thus confirming the need to examine such 3-word recurrences when investigating individual disciplinary discourses. The next most prominent feature is the significant numbers of LB that function as part of a numerative construction, many of which form extended numeratives used to measure. The nexus of these two features, disciplinary LB and numeratives, demonstrates a convention unique to this Discourse (Johns 2003). Numeratives are particularly useful for discussions of music dependent on music notation; more precisely, Western European staff notation. In fact, analysis of each lexical bundle Token and its ambient content revealed that no other systems of notation are connected to these numeratives. It is not that this is impossible — any form of music notation that represents music spatially could be discussed using such numeratives — but simply that it is not the case in the Expert Corpus. This in turn does not mean that other notational systems are never discussed in the Discourse, but rather that such discussion is exceedingly sparse as

compared with that based on staff notation. Furthermore, the ambient content accompanying the disciplinary-specific and numerative bundles confirms this dependence on score notation, revealing a preoccupation with the structural features of music as construed in a score, viewed as an artifact. Curiously, however, the even larger disciplinary concern is that of situating this music-as-artifact in its socio-historical context. While the practice of contextualization comes as no surprise, the substantial prioritization of it over the study of music-as-phenomenon is startling. Thus, the Discourse is centered on things rather than experiences. This is at once understandable — academic writing, with its penchant for nominalization, lends itself most readily to a focus on things — and regrettable because it ignores much of what has made and continues to make music attractive to every society: its performance and reception. By delimiting the Discourse thus, discussion of music is largely confined to a few facets of its production, while the various aspects of its performance and listeners' experience of it are ignored.

Among those facets left mostly unconsidered, there is a notable lack of reference to dynamics, articulation, and technique. This may point to a lack of nominal groups containing such disciplinary terms, yet such terminology is largely absent from the ambient content as

well, demonstrating that it is considered a less significant aspect of the Discourse and therefore of music generally. Perhaps dynamics, articulation, and technique are largely ignored because they tend, or are perceived, to operate at levels far below that of structure, and consequently are considered too elementary for discussion. If that is the case, however, it represents an undefended bias of the Discourse Community. Why should these facets of music be considered of less importance than structure, particularly given that they are frequently directly linked to the articulation of structure? Similarly, there is virtually no reference to silence or rests, despite the fact that all music emerges from, and merges again into, silence. Indeed, silence too is a fundamental means of articulating structure. As with the above facets, exclusion of silence from the discussion is yet another way in which the Discourse is constrained from engaging more fully with music-as-phenomenon. Of course, it could be argued that such aspects of music are relegated to score excerpts included as part of the multimodal Discourse. If that is the case, however, it would still indicate not only the Discourse's dependence on standard staff notation and variants thereof, but also a disinterest in discussing any such aspects.

Similar to expert writings, novice texts also feature disciplinary-specific lexical bundles and ambient content. Several prominent

numeratives and extended numeratives are included among these bundles. However, there is a noticeable discrepancy with the Expert Discourse in the types of disciplinary content found both in the bundles and ambient slots. Though some bundles are shared with the ExCo, many are not, particularly those that relate more to performance. There is also far less concern in the novice texts with socio-historical contextualization of pieces. Even the numeratives found in the NoCo just as often serve discussions of performances and recordings as of scores. Thus, novice writings are more performance-oriented, whereas experts center their discussions on notation (N.B.: the exception being Ethnomusicologists). There are references to various centuries in the novice texts, indicating that novices are being taught to write from a historical perspective; nonetheless, it is apparent from the discrepancy of focus within the two Corpora that novices are paradoxically expected to acquire the standards of the Discourse by responding to assignments that do not fully align with expert expectations. The fact that this occurs at both the Secondary and Tertiary level of novice production suggests that this paradox is imposed both by instructors who are performers and those who are Musicologists, as the former is more likely to teach coursework at the Secondary level, while the latter typically teaches disciplinary writing at the Tertiary. It would not necessarily be inappropriate to expect that novices write a somewhat different

genre than experts, provided this fact is explicitly identified to novices.

Given the above findings from both Corpora, there are three major narratives that arise from this research. Firstly, the Discourse claims to be a primary means of understanding musical meaning. It acts as an interpretive tool. This is especially noticeable in analysis and historical Musicology writings as experts claim to explicate original aspects of musical works through examination of notation, of the score. The fact that these writings tend to ignore certain fundamental aspects of music-as-phenomenon, however, indicates that certain elements of the music are sidelined in favor of others. This represents an incomplete approach to the study of music that could potentially alter the ongoing reception and re-presentation of musical pieces. In so doing it constitutes only a partial view of music. That is, the dependence on score notation (i.e., another semiotic), combined with the prestige afforded structure and harmony, significantly delimit the scope of the Discourse, effectively barring readers' access to many other facets of music and even types of music.

Secondly, the types of writing assigned to novices are not indicative of expert production; therefore, novices are inadvertently shut out

of the Discourse Community by the Community itself, despite the fact that members of it are the experts tasked with disciplinary writing instruction. This is implicit gate-keeping. Explicit gatekeeping arises from the Community's requirement that novices achieve fluency in the Discourse to be admitted. This challenge is complicated by the previous implicit practice. Thus, experts expect novices to achieve discourse competence with a type of writing that they do not provide opportunities to practice, which is unsurprising given the differing demands of writing assignments and articles. The essays that students are assigned are more often than not founded on specific scores or forms of notation, and therefore tend to ground students, — most of whom have come to the discipline by way of performance — in discussions of performance-related issues. No doubt, expert instructors perceive this focus on the score as a key feature of the Discourse, as indeed it is; nevertheless, the result is not a focus on the score so much as on other facets of the music and its performance. Without a frank admission on the part of the Discourse Community that it prizes the score above all else for the purposes of the Discourse, there can never be adequate assistance for novices to comprehend and acquire the Discourse norms. It lies, then, with the Community to objectively evaluate its actual practices and consider whether they should continue to thus underpin the Discourse, especially as it is precisely the reliance on staff notation

that continues to delimit its ability to break free of the Western canon to explore the entire range of music and musical experience. It seems very much that this problem, though pervasive, continues to go unrecognized; hence, Tenzer's (2015:23) statement once again that: 'Musicology, ethno- and otherwise, has pushed back hard…saying good riddance to the suffocating space of the fixed musical work, the biases of representation, the autonomy of musical sound'.

'It is a capital mistake to theorize before you have all the evidence. It biases the judgment.' (Doyle 1887:14). Holmes' dictum applies exactly to Tenzer's claim, which does not appear to be founded on any specific evidence. The third narrative to arise from this research, then, stands in sharp contrast to that claim. This study's findings demonstrate forcefully the extent to which Music Discourse remains dependent on score notation and therefore 'the fixed musical work' and 'the biases of representation'. Even more disturbing are the many specters conjured not just by the entire socio-historical context of music created through such notation, but especially those associated with the dominant timeframe stretching from the eighteenth to the early twentieth centuries. By focusing on music of that period, the Discourse is in danger of curating an elitist focus by prizing notated music of European and North American

composers of the eighteenth to early twentieth centuries, many of whom enjoyed the patronage of the aristocracy and (later) the bourgeois. These limitations are apparent in biases of gender (e.g., LB containing masculine pronouns, frequent mention of, or focus on, canonical male composers), of ethnicity (e.g., focus on the tradition and development of Western European staff notation), and geopolitical hegemony (e.g., LB containing 'United States' and 'New York'), all of which represent a form of inculcation of novices into these ideological norms. This is particularly alarming considering that the novices represented in this study all come from a distant culture with its own wealth of musical traditions. Of course, such associations are unavoidable to an extent, given the ways in which these inequities have shaped music throughout history. Nevertheless, it now lies within the power of the Discourse Community to redress these failings by recognizing the current state of the Discourse and subsequently taking steps to reorient and broaden it, giving space for every type of music and every facet of that music.

6.3 Recommendations

Accomplishing a reorientation of the Discourse would first necessitate moving away from the centrality of the score. This would certainly prove problematic, though, as the Discourse (taken in its most encompassing sense) is primarily concerned with

explaining the meaning of harmony and structure, a mode of analysis that is largely dependent on score notation. This is a fundamentally nineteenth-century concern, as partly evidenced by the dominant presence of temporal bundles containing 'Nineteenth century'. In fact, this disciplinary Discourse originated in mid-nineteenth century Germany as a means of constructing national identities through narration of their respective histories and cultures (Cohen 2014:185). Thus, it inevitably focused on the compositional giants of a given country's tradition. Perhaps the earliest generations of Musicologists saw discussion of form and structure as the best possible means of imbuing music with meaning in order to bolster that narrative. Whatever the case, the Discourse continues on a similar trajectory today in spite of the fact that experts claim a change of priorities.

As Ethnomusicology texts are least freighted with scores, they may serve as a point of departure for such change in the Discourse. Still, the cooperation of publishers, Music departments, and experts would be needed to work toward a creative confluence of discoursal forms that could serve a far wider and more diverse set of purposes. All of this would call for a good deal of creative effort, not the least of which should be the innovation and dissemination of materials that can properly guide novices (and experts) in the formulation of their writing and acquisition of the writing process and mechanics.

Regarding this last point, there are a few recommendations that can be made based on the present findings. Several of the non-disciplinary bundles could prove useful to novices, provided they are instructed how to use them accurately and cautioned not to overuse them. This would be true, for instance, of ExCo 1 'as well as', which can easily appear too frequently in a short essay. The same would apply to certain disciplinary bundles, particularly the extended 6- to 8-LB, as they tend not to appear more than once per text. Likewise, the most frequent LB could be taught, but with the caveat that those in the top 10% are likely overused, both the disciplinary and non-disciplinary ones alike. In a similar vein, low-frequency bundles from the Expert Discourse are used less often or well by novices, not unlike low-frequency terminology. Among these, ExCo 547 ('in this article I') could be of value to novice writers as it provides a typical frame for referring to one's own text (e.g., 'in this essay I') while simultaneously demonstrating that first-person singular pronouns are acceptable in the Discourse. Of course, the caveat to this is that the transitivity of the verbs succeeding this bundle would need to be studied to ensure that it was being used properly.

6.4 Limitations

Though this research has produced both substantial and significant findings, it is necessary to discuss its methodological limitations as

well. Because the methodology hinges on identification of formulaic items, it generates a view of Music Discourse through the lens of frequency. While this permits an overview of both recurrent lexical bundles and their ambient propositional content throughout the Expert Corpus, which in turn supports inferences about the broader Discourse, this does not permit a more granular view of the Discourse or its epistemological concerns. It is conceivable, therefore, that certain of those concerns may have been overlooked, those not directly connected to a lexical bundle. Were that additional propositional content to be mapped, as is the ambient content herein, it might reveal additional categories to be included in the cartographic analysis, or perhaps somewhat alter the proportions of the existing categories to one another. This might further be the case had the 2-word bundles been included in the study. As previously stated, however, analysis at such a granular level would have involved far more time and effort.

The type of overview offered herein also explicitly sets aside other types and instances of invention in this writing. This is not to imply that such invention should be devalued in any way. On the contrary, the unique contribution of each writer is a necessary and vital step in the process of building communal understanding. Nonetheless, instances of invention devoid of formulae have been omitted

precisely because this study aimed to produce a cartographic rather than an encyclopedic view of the Discourse. The latter could be useful but might also prove cumbersome for learners. A map makes salient information more readily available to novices and their instructors.

Another limitation is the selection of journals used for the Expert Corpus. A more comprehensive overview of the Discourse might have been gathered had articles been selected from, say, all of the top 50 or 100 ranking journals, which would thus include specialized journals. While this might create complications for representativeness, given such specializations, such a view could demonstrate which facets of music are trending in the Discourse. However, such a view would run the risk of better representing particular trends than the Discourse generally. Moreover, the abundance of so many additional texts would likely disproportionally favor non-disciplinary bundles, as the highest frequency LB in the current study are mostly grammatical in nature, with only a few containing generic disciplinary terms, such as 'music'.

Since the ambient content analyzed in this study is entirely nominal in function, it might also have been interesting to identify and categorize other types of content filling the ambient slots, such as

major and minor processes, or possibly even rank-shifted clauses. Still, this would have increased the workload considerably without yielding a sufficient amount of relevant information to justify the effort. As this is not a study of the grammar of Music Discourse, it was necessary to omit several grammatical functions from the analysis. Given their prevalence and interaction within texts, however, it would be interesting to analyze how lexical bundles serve a cohesive function within the Discourse. This was originally considered for the present study but ultimately abandoned as the sheer workload of the cartographic analysis became apparent.

Since corpora compiled according to differing criteria of representativeness could yield somewhat different findings, further investigation into the unique features of the sub-Discourses of Ethnomusicology and Music Theory could yield more detailed textual patterns. In the case of the former, it would be particularly interesting to see which 3-LB constitute disciplinary-specific bundles and how these may or may not be related to the findings of the present study. As for Music Theory, the present findings suggest that a narrower focus on that sub-Discourse would likely produce an even greater number of numerative bundles and more terminology-rich bundles.

As mentioned previously, the lack of multimodal analysis, capable of accounting for score excerpts and their interaction with the text, is an additional limitation; yet such work would not only once again have substantially increased the workload of this research, but also required it to move in a somewhat different direction than that suggested by frequency alone.

While the above limitations inevitably arose from the workload of the present methodology, the final two limitations were unavoidable for other reasons. The most obvious challenges to the verifiability and replicability of this study are the limited number of novice texts, including the sparse number of graduate student texts, of which there are only two master's student papers. The lack of doctoral-level writing creates a gap in the progression from Secondary novice to expert writing that interrupts an otherwise comprehensive view of the development of discourse competence. More problematic, however, is the convenience sampling encountered in the Novice Corpus. As stated previously, this was inescapable both because it was difficult to locate and contact potential contributors, they being few in number anyway, and because the collection period coincided with the onset of the COVID-19 pandemic and highly disruptive lock-down measures that were applied in its wake. Lack of a wider range of variability is problematic for representativeness (Biber

1993). To compensate for this dilemma, findings presented from the Novice Corpus have been appropriately hedged, as have comparisons of the two Corpora. As shown in the preceding chapter, though, much of the data from the NoCo falls into patterns predicted both by other studies and the results of the ExCo, suggesting that inferences based on its findings are generally reliable, even with its given its word count and number of texts.

6.5 Areas for Future Study

Having discussed limitations, it seems appropriate to mention several avenues for future research and publication suggested by the present study. Firstly, it would be instructive to conduct a comparative study of the findings from this research with those of a much larger corpus consisting of several articles each from the 50 or 100 most oft-cited journals. Were the results largely the same, this would offer powerful confirmation of the present findings regarding epistemological concerns in the Discourse. If the results were incongruent with the present findings, however, this might suggest that certain corners of the Discourse community have indeed moved beyond the canon as instantiated in score notation.

Secondly, there is certainly a need for further investigation of each sub-discipline represented here: Ethnomusicology, Musicology, and

Music Theory. As seen in the Global Views of complete texts (see pages 235-241), there are immediately observable differences in the use of high-frequency bundles in each of these sub-varieties of text. This then demonstrates a need to research each of these separately and in greater detail.

Thirdly, the participation of disciplinary-specific lexical bundles in this Discourse needs further analysis. Because these bundles include pronouns, numeratives, and terminology, they have the potential to serve a cohesive function within texts. If this is the case, such information could be valuable both to novices and expert instructors.

As mentioned above, the fourth recommendation is that future studies based on frequently recurring items be multimodal in nature. Given the limits of journal publication, it is likely that article score excerpts are of a relatively confined length, making them an object of study by the likely virtue of being approximately uniform in size.

Though admittedly tangential, the fifth recommendation is to focus research on the quotations found in this Discourse. Given the patterns discerned therein, it could be interesting to investigate the

degree to which such patterns originate with either the source material or the article's author, as this would indicate how such material is employed by the Discourse community to bolster its epistemological perspective.

Sixthly, it would surely prove fascinating to conduct a similar type of research with spoken texts, particularly those dialogues that constitute one-to-one performance lessons. The way in which performers attempt to convey musical ideas, information, and emotions to their students is surely fertile ground for linguistic research considering the fundamental impossibility of such a task. As Wittgenstein (1922:109) famously concluded, 'Wovon man nicht sprechen kann, darüber muß man schweigen.' Of course, teachers often do not have the luxury of remaining silent, so they resort to whatever means of signification they perceive necessary or available. This must result in a wealth of linguistic creativity, all of which has yet to be explored.

In addition to the above suggestions, there are two imminently practical applications of the present research. Firstly, a new textbook on Music writing is sorely needed. Such a book must include not just conventional instruction on sources and citation, but moreover a novice-friendly guide to constructing texts within the

sub-register of Music Discourse. This guide must include strategies for writing that are constructed from specific grammatical and discoursal advice, and also general information on the contextualization of propositional concerns within the Discourse and the normative epistemological concerns of the discipline.

Finally, and perhaps most challenging, is the recommendation to create a digital platform that would permit users to search key terms and their positions within lexical bundles, including 2-LB, from scores and possibly even texts. Insofar as this would require a corpus akin in scope to larger general corpora, it is something of a pipe dream, though no less worthy on account of that.

6.6 Conclusion

As explained in the introductory chapter, I undertook this research to satisfy my curiosity about a problem that had plagued me for years: How does Music Discourse work? What are its mechanics? As in most disciplines, experts within the Discourse Community write from their experience and intuition therefrom. Their skill renders any need to comprehend the mechanics unnecessary. This, combined with the numerous demands of teaching and research, likely lead many disciplinary writing instructors to rely on a few generic comments when providing feedback specifically about the

quality or mechanics of novice writing. In so doing, they inadvertently strand their students in a labyrinth of bewildering grammatical and discoursal choices. The novices whom I have taught bemoan the same questions and uncertainties I also faced as a novice: How do I begin? How can I make my writing clear? What does this feedback mean? How will I ever write ten pages! Hence, I designed this study not only to map Music Discourse by frequently recurrent lexical bundles and their ambient content, thus illuminating many of its typical structures and high-frequency items and propositional concerns, but most especially to teach myself the precise inner workings of this particular form of academic writing. That is to say, I did this to complete my own education as a disciplinary writing instructor. Aside from the numerous personal benefits this research has offered me, I hope that this work will be of service to novices and the experts who teach them, not only as instructors but also as keepers of the Discourse who bear the responsibility for transforming it into an all-encompassing discussion of music in all its manifestations. As the above findings demonstrate, the present state of the Discourse lags far behind the reality imagined for it by its practitioners. Rather than look to other experts for confirmation that the Community's aspirations are being achieved, this study has solely investigated the writings published by that Community for confirmation of the actual concerns and

biases of the Discourse. The results, having been derived empirically from two balanced and representative Corpora of expert and novice texts unequivocally lead to the conclusion that Music Discourse is delimited by the semiotic of score notation, a predisposition that arises from the prestige afforded music of the Western canon by the Discourse Community. Instead of permitting textual evidence to demonstrate this situation, experts turn to their experience for confirmation of their bias that, as Tenzer (2015:23) says, the Discourse has said 'good riddance to the suffocating space of the fixed musical work'. Gadamer's insight into language and preconceptions offers a word of caution for Tenzer's aspirational statement:

> One of the fundamental structures of all speaking is that we are guided by preconceptions and anticipations in our talking in such a way that these continually remain hidden and that it takes a disruption in oneself of the intended meaning of what one is saying to become conscious of these prejudices as such. (Gadamer 2004a:92)

End Notes

1. Platonism has provided one of the dominant approaches to the ontology of music. It posits that a single performance cannot contain the totality of all possible performances, nor even could the totality of all performances, rehearsals, scores, analyses and mental conceptions — were such a thing possible — be called the work (the absurdity of the argument notwithstanding) as these would still not include all such future potentials. While musical experiences may be regarded as historical events, musical works are understood as forms that manifest in tokens (Kivy 1983, 1987). This approach has been criticized (Goehr 1992; Davies, D. 2004), yet finding suitable alternatives has also proved vexing. Rohrbaugh (2003) offers an innovative solution that proposes art works be understood as *embodiments* of *historical individuals*. For an introduction to the foundational questions of music ontology, see Matheson and Caplan (2011). Sharpe (2004) provides a concise introduction with historical overview of the Philosophy of Music. Davies (2011) explores music and meaning. Lastly, for an ontology that excludes performance from its philosophical framework, read Cray's (2016) 'Unperformable Works and the Ontology of Music'.

2. I am indebted to Pope Paul VI's (Paulus PP. VI 1965) formulation of *re-presentation,* in his encyclical on the nature of the Mass, for

the inspiration to employ the notion of *making present*. Though not employing the substance of his theological argument — I argue that such re-presentation brings the music once more before us, rather than placing us before the musical fact — I do appropriate the notion that a historical event may be re-presented in other times by other agents, and be at once a temporal manifestation of a supra-temporal existent. The original formulation reads: 'Tunc Dominus incruente immolatur in Sacrificio Missae, Crucis sacrificium repraesentante et virtutem eius salutiferam applicante, cum per consecrationis verba sacramentaliter incipit praesens adesse, tamquam spiritualis fidelium alimonia, sub speciebus panis et vini'.

3. Skehan notes that a large number of applied linguists, himself included, hold first degrees in seemingly unrelated subjects (Skehan 2016).

4. The United States Central Intelligence Agency's statistical apparatus estimates the population of China (inclusive of Taiwan, Hong Kong and Macau) as 1,379,302,771 as of July 2017, making it the world's most populous nation. Economically, the CIA also ranked China first globally as of 2017, using purchasing power parity (PPP) to adjust for GDP measured at market exchange rates (Central Intelligence Agency 2018), and their statistics are confirmed by The World Bank (2018). For further

explanation of this argument, see Smith (2017), 'Who Has the World's No. 1 Economy? Not the U.S.'.

5. On the rise of Western classical music in China, see Brahmstedt & Brahmstedt (1997), 'Music Education in China'; Law and Ho (2011), 'Music Education in China'; Huang (2012), 'Why Chinese people play Western classical music'; Kahn & Wakin (2007), 'Classical Music Looks Toward China With Hope'; Montefiore (2014), 'Why piano-mania grips China's children'; and for information relating to Ningbo, read Hernández's (2016) 'Steinway's Grand Ambitions for Its Pianos in China'. For the historical development of this phenomena, see Melvin and Cai's (2004) *Rhapsody in Red: How Western Classical Music Became Chinese*.

6. According to the Chinese Ministry of Education, 4,040,210 Chinese nationals were studying abroad in 2015. Approximately 45% of those students returned to China upon graduation. Of those two million plus, a significant percent studied in the United States, Great Britain, Australia and Canada (44.99% of doctoral candidates, 30.43% of master's students, and 37.59% of undergraduates) [N.B.: The following information does not appear to be provided in English translation by the Ministry.]《博士研究生学历的留学回国人员主要留学的国家为美国（28.95%）、日本（12.90%）、英国（9.95%）、法国（6.48%）、德国（5.57%）、韩

国（4.91%）、新加坡（4.58%）、澳大利亚（3.60%）和加拿大（2.49%）。硕士研究生学历的留学回国人员留学英国的人数最多，占42.52%，其次是美国（18.83%）、澳大利亚（10.41%）、法国（4.24%）、韩国（3.26%）、日本（3.04%）、德国（2.40%）、俄罗斯（1.65%）和加拿大（1.19%）。本科和专科学历的留学回国就业人员中，21.27%留学韩国，其次为英国（10.63%）、美国（9.56%）、澳大利亚（8.86%）、加拿大（8.54%）、日本（7.97%）、俄罗斯（7.72%）、马来西亚（4.49%）、新西兰（3.10%）和乌克兰（1.90%）。》。中华人民共和国教育部（2018）《中国留学回国就业蓝皮书2015》情况介绍. Available at: http://www.moe.edu.cn/jyb_xwfb/xw_fbh/moe_2069/xwfbh_2016n/xwfb_160325_01/160325_sfcl01/201603/t20160325_235214.html (accessed 16 March 2018).]

References:

Thesis

Ädel, A., & Erman, B. (2012). 'Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach'. *English for Specific Purposes*, *31*(2), 81–92. https://doi.org/10.1016/j.esp.2011.08.004

Alhassan, L., & Wood, D. (2015). 'The effectiveness of focused instruction of formulaic sequences in augmenting L2 learners' academic writing skills: A quantitative research study'. *Journal of English for Academic Purposes*, *17*, 51–62. https://doi.org/10.1016/j.jeap.2015.02.001

Altenberg, B. (1998). 'On the phraseology of spoken English: the evidence of recurrent word-combinations'. In Cox, K. (ed.) *EAP now! : English for academic purposes* (2nd ed., pp. 101–122). Frenchs Forest, NSW : Pearson Longman.

Altenberg, B., & Granger, S. (2001). 'The Grammatical and Lexical Patterning of MAKE in Native and Non-Native Student Writing'.

Applied Linguistics, *22*(2), 173–194. https://doi.org/10.1093/applin/22.2.173

Anthony, L. (2017). AntFileConverter (Version 1.2.1) [Macintosh OS X]. Tokyo, Japan: Waseda University. Available at: http://www.laurenceanthony.net/software

Anthony, L. (2018). AntConc (Version 3.5.3) [Macintosh OS X]. Tokyo, Japan: Waseda University. Available at: http://www.laurenceanthony.net/software

Antoch, J., Prchal, L., & Sarda, P. (2013). 'Combining Association Measures for Collocation Extraction Using Clustering of Receiver Operating Characteristic Curves'. *Journal of Classification*, *30*(1), 100–123. https://doi.org/10.1007/s00357-013-9123-x

Atkinson, D. (1999). 'LANGUAGE AND SCIENCE'. *Annual Review of Applied Linguistics*, *19*, 193–214. https://doi.org/10.1017/S026719059919010X

Aull, L. L., & Lancaster, Z. (2014). 'Linguistic Markers of Stance in Early and Advanced Academic Writing'. *Written*

Communication, *31*(2), 151–183. https://doi.org/10.1177/0741088314527055

Baker, P. (2006). *Using corpora in discourse analysis.* London ; New York: Continuum.

Barker, A. (ed.). (1984). *Greek Musical Writings, I: The Musician and his Art*. Cambridge: Cambridge University Press.

Barker, A. (ed.). (1989). *Greek Musical Writings, II: Harmonic and Acoustic Theory*. Cambridge: Cambridge University Press.

Barton, E. (1993). 'Evidentials, Argumentation, and Epistemological Stance'. *College English*, *55*(7), 745–769. https://doi.org/10.2307/378428

Basturkmen, H. (2010). *Developing courses in English for specific purposes* [electronic resource]. Basingstoke: Palgrave Macmillan.

Basturkmen, H., & Wette, R. (2016). 'English for academic purposes', in Hall, G. (ed.), *The Routledge Handbook of*

English language teaching (pp. 164–176). Milton Park, Abingdon, Oxon ; New York, NY: Routledge.

Belcher, D. D. (2009). *English for specific purposes in theory and practice*. Ann Arbor: University of Michigan Press.

Bellman, J. D. (2006). *A Short Guide to Writing About Music* (2nd ed.). London: Pearson.

Berg, J. L. (2015). *A Learner Needs Analysis for IB Music.* MA Dissertation, University of Nottingham Ningbo, China.

Bestgen, Y. (2017). 'Beyond single-word measures: L2 writing assessment, lexical richness and formulaic competence'. *System*, *69*, 65–78.

Bhatia, V. K. (2004). *Worlds of written discourse.* London ; New York: Continuum.

Biber, D., 1990. Methodological issues regarding corpus-based analyses of linguistic variation. Literary and Linguistic Computing 5, 257–269.

Biber, D. (1993). 'Representativeness in Corpus Design'. *Literary and Linguistic Computing*, *8*(4), 243–257.

Biber, D. (2010). 'What can a corpus tell us about registers and genres?', in *The Routledge handbook of corpus linguistics*, O'Keeffe, A. & McCarthy, M. (eds.), (pp. 241–254). London: Routledge.

Biber, D., & Barbieri, F. (2007). 'Lexical Bundles in University Spoken and Written Registers'. *English for Specific Purposes*, *26*(3), 263–286. https://doi.org/10.1016/j.esp.2006.08.003

Biber, D., Connor, U., & Upton, T.A. (2007). *Discourse on the move: using corpus analysis to describe discourse structure.* Amsterdam; Philadelphia: John Benjamins Publishing Company.

Biber, D., & Conrad, S. (2002). 'Quantitative Corpus-based Research: Much More than Bean Counting.' *TESOL Quarterly*, *35*(2), 331–36.

Biber, D., & Conrad, S. (2009). *Register, genre, and style.* Cambridge, UK ; New York: Cambridge University Press.

Biber, D., & Conrad, S. (2015). 'Register Variation: A Corpus Approach', in *The handbook of discourse analysis,* (2nd ed.), Tannen, D., Hamilton, H. E., Schiffrin, D., & Adger, C. T. (eds.), (pp. 505–525). Chichester, England: Wiley Blackwell.

Biber, D., Conrad, S., & Cortes, V. (2004). 'If you look at...: Lexical bundles in university teaching and textbooks'. *Applied Linguistics*, *25*(3), 371–405. https://doi.org/10.1093/applin/25.3.371

Biber, D., Conrad, S., & Reppen, R. (1996). 'Corpus-Based Investigations of Language Use'. *Annual Review of Applied Linguistics*, *16*, 115–136. https://doi.org/10.1017/S0267190500001471

Biber, D., & Gray, B. (2010). 'Challenging Stereotypes about Academic Writing: Complexity, Elaboration, Explicitness'. *Journal of English for Academic Purposes*, *9*(1), 2–20. https://doi.org/10.1016/j.jeap.2010.01.001

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English.* Harlow, England ; New York : Longman.

Block, D. (2003). *The social turn in second language acquisition.* Washington, DC: Georgetown University Press.

Blommaert, J. (2005). *Discourse: a critical introduction.* New York: Cambridge University Press.

Bohlman, P. V. (1999). 'Rethinking "Thinking Music"', in *Rethinking music,* Cook, N. & Everist, M. (eds.), (pp. 17–34). Oxford; New York: Oxford University Press.

Bondi, M. (2008). 'Emphatics in academic discourse: Integrating corpus and discourse tools in the study of cross-disciplinary variation', in *Corpora and discourse: the challenges of different settings,* Ädel, A. & Reppen, R. (eds.), (pp. 31–55). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Bondi, M. (2015). '*Probably most important of all:* Importance markers in academic and popular history articles', in *Corpora,*

grammar and discourse: in honour of Susan Hunston, Groom, N., Charles, M., John, S., & Hunston, S. (eds.), (pp.161-182). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Bordet, G. (2015). 'The Role of "Lexical Paving" in Building a Text According to the Requirements of a Target Genre', in *English for academic purposes: approaches and implications*, Paul, T. & Diani, G. (eds.), (pp.43-77). Available at: https://ebookcentral.proquest.com/lib/nottingham/detail.action?docID=4534871

Bowker, L. & Pearson, J. 2002. Working with specialized language [electronic resource] : a practical guide to using corpora/ Lynne Bowker and Jennifer Pearson. London ; New York : Routledge, London ; New York.

Bowman, W. D. (1998). *Philosophical perspectives on music.* New York: Oxford University Press.

Brahmstedt, H., & Brahmstedt, P. (1998). 'Music Education in China'. *Music Educators Journal*, *83*(6), 28–30,52. https://doi.org/10.2307/3399021

Brown, D. E. (1991). *Human Universals*. New York: McGraw-Hill.

Brown, G., & Yule, G. (2000). *Discourse analysis.* Beijing: Foreign Language Press.

Bruce, I. (2008). *Academic writing and genre: a systematic analysis*. London: Continuum.

Bruce, I. (2011). *Theory and concepts of English for academic purposes.* Basingstoke: Palgrave Macmillan.

Brumfit, C. J. (1995). 'Teacher professionalism and research', in *Principle & practice in applied linguistics: studies in honour of H. G. Widdowson,* Cook, G. & Seidlhofer, B. (eds.), (pp. 27–41). Oxford: Oxford University Press.

Butt, D., Fahey, R., Feez, S., Spinks, S. & Yallop, C. (2000). *Using functional grammar: an explorer's guide,* (2nd ed.). Sydney, NSW: National Centre for English Language Teaching and Research.

Callies, M. (2015). 'Learner corpus methodology', in *The Cambridge handbook of learner corpus research,* Granger, S., Gilquin, G. & Meunier, F. (eds.), (pp. 35–55). Cambridge: Cambridge University Press.

Campbell, A. T. (2006). 'Malinowski, Bronislaw Kaspar (1884-1942)', in *Encyclopedia of language & linguistics,* (2nd ed.), [electronic resource], Brown, E. K. (ed.), (pp. 467–468). Amsterdam: Elsevier. Available at: http://www.sciencedirect.com/science/referenceworks/9780080448541

Canagarajah, A. S. (1999). *Resisting linguistic imperialism in English teaching.* Oxford: Oxford University Press.

Carter, R., McCarthy, M., 1995. Grammar and the spoken language. Applied Linguistics 16, 141–158.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English: a comprehensive guide: spoken and written English grammar and usage.* Cambridge; New York: Cambridge University Press.

Carver, D. (1983). 'Some propositions about ESP'. *The ESP Journal*, *2*(2), 131–137. https://doi.org/10.1016/0272-2380(93)90003-P

Central Intelligence Agency. (2018) 'Publishing Domain'. *The World Factbook*. Available at: https://www.cia.gov/library/publications/the-world-factbook/geos/ch.html [Accessed 16 March, 2018].

Chang, P. (2012). 'Using a stance corpus to learn about effective authorial stance-taking: a textlinguistic approach'. *ReCALL*, *24*(2), 209–236. https://doi.org/10.1017/S0958344012000079

Charles, M., Pecorari, D., & Hunston, S. (2009). *Academic writing: at the interface of corpus and discourse* [electronic resource]. London ; New York : Continuum.

Chen, Y. H., & Baker, P. (2010). 'Lexical Bundles in L1 and L2 Academic Writing'. *Language Learning & Technology*, *14*(2), 30–49.

Chen, Y.-H., & Baker, P. (2016). 'Investigating Criterial Discourse Features across Second Language Development: Lexical Bundles in Rated Learner Essays, CEFR B1, B2 and C1'. *Applied Linguistics*, *37*(6), 849–880. https://doi.org/10.1093/applin/amu065

Cheng, W. (2012). *Exploring corpus linguistics: language in action* [electronic resource]. Abingdon, Oxon: Routledge.

Cho, S. (2004). 'Challenges of Entering Discourse Communities Through Publishing in English: Perspectives of Nonnative-Speaking Doctoral Students in the United States of America.' *Journal of Language, Identity & Education*, *3*(1), 47–72. https://doi.org/10.1207/s15327701jlie0301_3

Coffin, C., & Hewings, A. (2003). 'Writing for different disciplines', in *Teaching academic writing: a toolkit for higher education,* (pp. 45–72). London: Routledge. Available at: *http://www.myilibrary.com?id=11471*

Cohen, B. 2014. 'Limits of National History: Yoko Ono, Stefan Wolpe, and Dilemmas of Cosmopolitanism.' *The Musical*

Quarterly 97, 181–237. https://doi.org/10.1093/musqtl/gdu008

Conrad, S. (2002). 'Corpus Linguistic Approaches for Discourse Analysis'. *Annual Review of Applied Linguistics*, *22*. https://doi.org/10.1017/S0267190502000041

Corrigan, R. (2009). *Formulaic language.* Philadelphia: John Benjamins.

Corrigan, R., Moravcsik, E., Ouali, H., & Wheatley, K. (2009). 'Introduction: Approaches to the study of formulae', in *Formulaic language: Volume 1, Distribution and historical change* [electronic resource]*,* Corrigan, R., Moravcsik, E., Ouali, H., & Wheatley, K. (eds.)*,* (pp. XI–XXIV). Philadelphia: John Benjamins. Available at: http://site.ebrary.com/lib/unnc/Doc?id=10310858

Cortes, V. (2002). 'Lexical bundles in Freshman composition', in *Using corpora to explore linguistic variation,* Reppen, R., Fitzmaurice, S. M. & Biber, D. (eds.), (pp. 131–145). Amsterdam: John Benjamins Publishing Company.

Cortes, V. (2004). 'Lexical bundles in published and student disciplinary writing: Examples from history and biology'. *English for Specific Purposes*, *23*(4), 397–423. https://doi.org/10.1016/j.esp.2003.12.001

Cortes, V, Csomay, E., & Biber, D. (2015). *Corpus-based research in applied linguistics: studies in honor of Doug Biber.* Amsterdam; Philadelphia: John Benjamins Publishing Company.

Crawford, W. J., & Csomay, E. (2016). *Doing corpus linguistics.* New York: Routledge.

Cray, W. D. (2016). 'Unperformable Works and the Ontology of Music'. *British Journal of Aesthetics*, *56*(1), 67–81. https://doi.org/10.1093/aesthj/ayv047

Crossley, S. A., Cai, Z., & McNamara, D. S. (2012). 'Syntagmatic, Paradigmatic, and Automatic N-gram Approaches to Assessing Essay Quality', in *Proceedings of the Twenty-Fifth International Florida Artificial Intelligence Research Society Conference* (pp. 214–219). Association for the Advancement of Artificial Intelligence. Available at: www.aaai.org.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). 'The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality'. *Journal of Second Language Writing*, *32*(Supplement C), 1–16. https://doi.org/10.1016/j.jslw.2016.01.003

Crosthwaite, P., & Jiang, K. (2017). 'Does EAP affect written L2 academic stance? A longitudinal learner corpus study'. *System*, *69*, 92–107. https://doi.org/10.1016/j.system.2017.06.010

Cummings, M. (2006). 'Halliday, Michael A. K. (b. 1925)', in *Encyclopedia of language & linguistics* (2nd ed.), [electronic resource], Brown, E. K. (ed.), (pp. 196–197). Amsterdam: Elsevier. Available at: http://www.sciencedirect.com/science/referenceworks/9780080448541

Crystal, D. (2010). *The Cambridge encyclopedia of language* (3rd ed.). Cambridge: Cambridge University Press.

Cunningham, K. J. (2017). 'A phraseological exploration of recent mathematics research articles through key phrase frames'. *Journal of English for Academic Purposes*, *25*, 71–83.

Davies, D. (2004). *Art as Performance*. Malden: Blackwell.

Davies, S. (2011). *Musical understandings: and other essays on the philosophy of music* [electronic resource]. Oxford: Oxford University Press. Available at: http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199608775.001.0001/acprof-9780199608775

De Chazal, E. (2014). *English for academic purposes.* Oxford: Oxford University Press.

De Cock, S. (1998). 'A recurrent word combination approach to the study of formulae in the speech of native and non-native speakers of English'. *International Journal of Corpus Linguistics*, *3*(1), 59–80.

De Cock, S., Granger, S., Leech, G., & McEnery, T. (1998). 'An automated approach to the phrasicon of EFL learners', in

Learner English on computer, Granger, S. (ed.), (pp. 67–79). London: Longman.

Dewaele, J.-M. (2018). 'Why the Dichotomy "L1 Versus LX User" is Better than "Native Versus Non-native Speaker."' *Applied Linguistics*, *39*(2), 236–240.

Díez-Bedmar, M. B., & Casas Pedrosa, A. V. (2011). 'The use of prepositions by Spanish learners of English at university level: A longitudinal analysis', in *Corpora, Language, Teaching, and Resources: From Theory to Practice,* Kübler, N. (ed.), (pp. 199–218). Bern: Peter Lang.

Doughty, C., & Long, M. H. (2003). 'The Scope of Inquiry and Goals of SLA', in *The handbook of second language acquisition,* Doughty, C., & Long, M. H. (eds.), (pp. 3–16). Malden, Mass; Oxford: Blackwell.

Doyle, A. C. (1887). 'A Study in Scarlet', in *Beeton's Christmas Annual,* (pp.). London, New York, Melbourne: Ward Lock & Co. Accessed on 05-04-2024 from: https://www.fulltextarchive.com/book/A-Study-In-Scarlet/

Duckles, V., Pasler, J., Stanley, G., Christensen, T., Haggh, B. H., Balchin, R., Libin, L., Seebass, T., Page, J. K., Goehr, L., Bujic, B., Clarke, E. F., McClary, S., Gribenski, J., Gianturco, C., Potter, P. M., Fallows, D., Velimirović, M., Tomlinson, G., Béhague, G., Kanazawa, M., & Platt, P. (2001). 'Musicology', in *Grove Music Online*. Available at: http://www.oxfordmusiconline.com/grovemusic/view/10.1093/gmo/9781561592630.001.0001/omo-9781561592630-e-0000046710?rskey=o7hTvh&result=2

Dudley-Evans, T., & St. John, M. J. (1998). *Developments in ESP: a multi-disciplinary approach.* Cambridge, U.K.; New York: Cambridge University Press.

Durrant, P. (2009). 'Investigating the Viability of a Collocation List for Students of English for Academic Purposes'. *English for Specific Purposes*, *28*(3), 157–169. https://doi.org/10.1016/j.esp.2009.02.002

Dutra, D.P., Orfanó, B.M. & Almeida, V.C. 2019. Result linking adverbials in learner corpora. Domínios de Lingu@gem 13, 400–431. https://doi.org/10.14393/DL37-v13n1a2019-17

Ebeling, S. O., & Hasselgård, H. (2015). 'Learner corpora and phraseology', in *The Cambridge handbook of learner corpus research,* Granger, S., Gilquin, G., & Meunier, F. (eds.), (pp. 207–229). Cambridge: Cambridge University Press.

Eegermann, H., Efernando, N., Echuen, L., & Emcadams, S. (2015). 'Music Induces Universal Emotion-Related Psychophysiological Responses: Comparing Canadian Listeners To Congolese Pygmies'. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01341

Eggins, S. (2004). *An introduction to systemic functional linguistics* (2nd ed.). New York: Continuum.

Ellis, N. C. (2003). 'Constructions, Chunking, and Connectionism: The Emergence of Second Language Structure', in *The handbook of second language acquisition,* Doughty, C., & Long, M. H. (eds.), (pp. 64–103). Malden, Mass; Oxford: Blackwell.

Ellis, N. C. (2012). 'Formulaic Language and Second Language Acquisition: Zipf and the Phrasal Teddy Bear'. *Annual Review*

of Applied Linguistics, *32*, 17–44. https://doi.org/10.1017/S0267190512000025

Ellis, N. C. (2014). 'Frequency-based accounts of second language acquisition', in *The Routledge handbook of second language acquisition,* Gass, S. M. & Mackey, A. (eds.), (pp. 193-210). London: Routledge, Taylor & Francis Group.

Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford; New York: Oxford University Press.

Ellis, R. (2015). *Understanding second language acquisition* (2nd ed.). Oxford: Oxford University Press.

Englebretson, R. (2007). *Stancetaking in discourse: subjectivity, evaluation, interaction* [electronic resource]. Amsterdam; Philadelphia: John Benjamins Publishing Company.

English, F., & Marr, T. (2015). *Why do linguistics?: reflective linguistics and the study of language*. London; New York: Bloomsbury Academic.

Escudero, L. M. (2011). 'A corpus-based insight into genre: The case of WIPO domain name arbitration decisions'. *Discourse & Communication*, *5*(4), 375–392. https://doi.org/10.1177/1750481311427084

Evison, J. (2010). 'What are the basics of analysing a corpus?', in *The Routledge handbook of corpus linguistics,* O'Keeffe, A. & McCarthy, M. (eds.), (pp. 122–135). London: Routledge.

Faigley, L., & Hansen, K. (1985). 'Learning to Write in the Social Sciences'. *College Composition and Communication*, *36*(2), 140–49. https://doi.org/10.2307/357434

Fetzer, A. (2012). 'Textual coherence as a pragmatic phenomenon', in *The Cambridge handbook of pragmatics,* Allan, K. & Jaszczolt, K. M. (eds.)*.* Cambridge: Cambridge University Press.

Firth, J. R. (1969) [1935]. 'The Technique of Semantics', in *Papers in Linguistics: 1934-1951* (pp. 7-33). London: Oxford University Press.

Firth, J. R. (1957). *A synopsis of linguistic theory*. Oxford: Blackwell.

376 of 460

Flowerdew, J. (1999a). 'Writing for scholarly publication in English: The case of Hong Kong.' *Journal of Second Language Writing*, *8*(2), 123–145. https://doi.org/10.1016/S1060-3743(99)80125-8

Flowerdew, J. (1999b). 'Problems in writing for scholarly publication in English: The case of Hong Kong.' *Journal of Second Language Writing*, *8*(3), 243–264. https://doi.org/10.1016/S1060-3743(99)80116-7

Flowerdew, J. (2015c). 'Some thoughts on English for Research Publication Purposes (ERPP) and related issues', *48*(2), 250–262. https://doi.org/10.1017/S0261444812000523

Flowerdew, J. (2016). 'Introduction: Discourse in context', in *Discourse in context,* (Paperback ed.), Flowerdew, J. (ed.), (pp. 1–25). London: Bloomsbury Academic.

Flowerdew, J., & Peacock, M. (2001). *Research perspectives on English for academic purposes.* Cambridge: Cambridge University Press.

Flowerdew, L. (2002). 'Corpus-Based Analyses in EAP', in *Academic discourse* [electronic resource], Flowerdew, J. (ed.), (pp. 95–114). Harlow: Longman. Available at: http://nottingham.eblib.com/patron/FullRecord.aspx?p=1710583

Flowerdew, L. (2004). 'The Argument for Using English Specialized Corpora to Understand Academic and Professional Settings', in *Discourse in the professions: perspectives from corpus linguistics* [electronic resource], Connor, U. & Upton, T. A. (eds.), (pp. 11–33). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Flowerdew, L. (2014). 'Which Unit for Linguistic Analysis of ESP Corpora of Written Text?', in *Corpus analysis for descriptive and pedagogical purposes: ESP perspectives,* Gotti, M. & Giannoni, D. S. (eds.), (pp. 25–41). Bern: Peter Lang.

Foucault, M.. (2002a). *The archaeology of knowledge,* Smith, A. M. S. (trans.). London: Routledge.

Foucault, M. (2002b). *The order of things: an archaeology of the human sciences,* Tavistock/Routledge (trans.)*.* London: Routledge.

Fontaine, L., Bartlett, T., & O'Grady, G. (eds.) (2013). *Systemic functional linguistics: exploring choice.* Cambridge, United Kingdom; New York: Cambridge University Press.

Frankenberg-Garcia, A. (2016). 'Corpora in ELT', in *The Routledge Handbook of English language teaching,* Hall, G. (ed.), (pp. 383–398). Milton Park, Abingdon, Oxon; New York, NY: Routledge.

Friginal, E., & Hardy, J. A. (2014). *Corpus-based sociolinguistics: a guide for students.* New York ; London: Routledge.

Gadamer, H. G. (2004a). *Philosophical Hermeneutics.* Linge, D. E. (trans. & ed.). Berkeley, Los Angeles, London: University of California Press.

Gadamer, H. G. (2004b). *Truth and method.* (2nd rev. ed.), Weinsheimer, J., & Marshall, D. G. (trans.). London, New York: Continuum.

Gebhard, M., Chen, I.-A., Graham, H., & Gunawan, W. (2013). 'Teaching to mean, writing to mean: SFL, L2 literacy, and

teacher education'. *Journal of Second Language Writing*, *22*(2), 107–124. https://doi.org/10.1016/j.jslw.2013.03.005

Gee, J. P. (1992). *The Social Mind: Language, Ideology, and Social Practice*. New York: Bergin & Garvey.

Gee, J. P. (1999). 'The New Literacy Studies and the "Social Turn."' ERIC: Institute of Education Sciences. Available at: https://eric.ed.gov/?id=ED442118

Gee, J. P. (2014a). *An introduction to discourse analysis: theory and method,* (4th ed.). London: Routledge.

Gee, J. P. (2014b). 'Tools of Inquiry and Discourses', in *The discourse reader* (3rd ed.), Jaworski, A., & Coupland, N. (eds.), (pp. 142–153). London: Routledge.

Gesuato, S., 2011. Structure, content and functions of calls for conference abstracts, in: Bhatia, V., Sánchez Hernández, P., Pérez-Paredes, P. (Eds.), Researching Specialized Languages. John Benjamins Publishing Company, Amsterdam, NETHERLANDS, pp. 47–70.

Gilquin, G. (2015). 'From design to collection of learner corpora', in *The Cambridge handbook of learner corpus research,* Granger, S., Gilquin, G., & Meunier, F. (eds.), (pp. 9–34). Cambridge: Cambridge University Press.

Gimenez, J. (2011). 'Disciplinary epistemologies, generic attributes and undergraduate academic writing in nursing and midwifery'. *Higher Education*, *63*(4), 401–419. https://doi.org/10.1007/s10734-011-9447-6

Gledhill, C. (2015). 'On the Phraseology of Grammatical Items in Lexico-Grammatical Patterns and Science Writing', in *English for academic purposes: approaches and implications*, Paul, T. & Diani, G. (eds.), (pp.11-41). Available at: https://ebookcentral.proquest.com/lib/nottingham/detail.action?docID=4534871

Goehr, L. (1992). *The imaginary museum of musical works: an essay in the philosophy of music.* Oxford: Clarendon Press.

Graesser, A. C., Gernsbacher, M. A., & Goldman, S. R. (2003). 'Introduction to the Handbook of Discourse Processes', in *Handbook of discourse processes,* Graesser, A. C.,

Gernsbacher, M. A., & Goldman, S. R. (eds.), (pp. 1–23). New York: Routledge.

Granger, S. (2009). 'The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation', in *Corpora and language teaching* Aijmer, K. (ed.)*. Amsterdam; Philadelphia: John Benjamins Publishing Company.

Granger, S., Gilquin, G., & Meunier, F. (2015). 'Introduction: learner corpus research — past, present and future', in *The Cambridge handbook of learner corpus research,* Granger, S., Gilquin, G., & Meunier, F. (eds.), (pp. 1–5). Cambridge: Cambridge University Press.

Green, C. F., Christopher, E. R., & Mei, J. L. K. (2000). 'The Incidence and Effects on Coherence of Marked Themes in Interlanguage Texts: A Corpus-Based Inquiry'. *English for Specific Purposes*, *19*(2), 99–113.

Grimshaw, A. D. (2003). 'Genres, Registers, and Contexts of Discourse', in *Handbook of discourse processes,* Graesser, A.

C., Gernsbacher, M. A., & Goldman, S. R. (eds.), (pp. 25–82). New York: Routledge.

Groom, N., Charles, M., John, S., & Hunston, S. (2015). 'Corpora, grammar, and discourse analysis: Recent trends, current challenges', in *Corpora, grammar and discourse: in honour of Susan Hunston,* Groom, N., Charles, M., John, S., & Hunston, S. (eds.), (pp.1-19). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Hall, G. (2016). 'Introduction: English language teaching in the contemporary world', in *The Routledge Handbook of English language teaching,* Hall, G. (ed.), (pp. 1–10). Milton Park, Abingdon, Oxon ; New York: Routledge.

Halliday, M. A. K. (2001). *Language as social semiotic: the social interpretation of language and meaning.* Beijing: Foreign Language Press.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.), [electronic resource]. Abingdon, Oxon.: Routledge.

Halliday, M. A. K., McIntosh, A., & Strevens, P. (1964). *The linguistics sciences and language teaching*. London: Longman.

Harmony. (2024). Britannica Academic. Retrieved: 22 April 2024, from https://academic-eb-com.proxy.lib.duke.edu/levels/collegiate/article/harmony/110125#article-contributors

Hatami, S. (2015). 'Teaching Formulaic Sequences in the ESL Classroom'. *TESOL Journal*, *6*, 112–129. https://doi.org/10.1002/tesj.143

Herbert, T. (2009). *Music in Words: A Guide to Researching and Writing about Music*. Oxford: Oxford University Press.

Hernández, J. C. (2016, July 9). 'Steinway's Grand Ambitions for Its Pianos in China'. *The New York Times*. Available at: https://www.nytimes.com/2016/07/10/business/international/steinways-grand-ambitions-in-china.html

Hiltunen, T., & Mäkinen, M. (2014). 'Formulaic Language in Economics Papers: Comparing Novice and Published Writing', in *Corpus analysis for descriptive and pedagogical purposes:*

ESP perspectives, Gotti, M. & Giannoni, D. S. (eds.), (pp. 345–368). Bern: Peter Lang.

Hitchcock, H.W., Deaville, J., 2013. Musicology. Oxford University Press.

Hoey, M. (2001). *Textual interaction: an introduction to written discourse analysis.* London: Routledge.

Holoman, D. K. (2014). *Writing about Music: A Style Sheet* (Third Edition). Oakland, California: University of California Press.

Howarth, P. (1998). 'Phraseology and Second Language Proficiency'. *Applied Linguistics*, *19*(1), 24–44. https://doi.org/10.1093/applin/19.1.24

Howatt, A.P.R. & Widdowson, H. G. (2004). *A history of English language teaching* (2nd ed.). Oxford: Oxford University Press.

Huang, H. (2012). 'Why Chinese people play Western classical music: Transcultural roots of music philosophy'. *International Journal of Music Education*, *30*(2), 161–176. https://doi.org/10.1177/0255761411420955

Hunston, S. (2002). *Corpora in applied linguistics.* Cambridge: Cambridge University Press.

Hunston, S., & Francis, G. (2000). *Pattern grammar: a corpus-driven approach to the lexical grammar of English.* Amsterdam; Philadelphia: John Benjamins Publishing Company.

Hutchinson, T., & Waters, A. (1987). *English for specific purposes: a learning-centred approach.* Cambridge: Cambridge University Press.

Hyland, K. (1994). 'Hedging in academic writing and EAP textbooks'. *English for Specific Purposes*, *13*(3), 239–256. https://doi.org/10.1016/0889-4906(94)90004-3

Hyland, K. (2006). *English for academic purposes: an advanced resource book.* London: Routledge.

Hyland, K. (2008). 'Academic clusters: text patterning in published and postgraduate writing'. *International Journal of Applied*

Linguistics, *18*(1), 41–62. https://doi.org/10.1111/
j.1473-4192.2008.00178.x

Hyland, K. (2009). *Academic discourse: English in a global context.*
London; New York: Continuum.

Hyland, K. (2012). *Disciplinary identities: individuality and
community in academic writing.* Cambridge: Cambridge
University Press.

International Baccalaureate®. (2018b). Extended essay | Diploma.
Available at: http://www.ibo.org/about-the-ib/ [Accessed 21
March, 2018].

Irvine, D., & Radice, M. (2003). *Irvine's Writing About Music: Third
Edition* (Third Edition). Milwaukee, Wisconsin: Amadeus Press.

Jacoby, S., & Gonzalez, P. (1991). 'The constitution of the expert-
novice in scientific discourse'. *Issues in Applied Linguistics*,
2(2), 149–181.

Jacoby, S., & Ochs, E. (1995). 'Co-construction: an introduction'. *Research on Language and Social Interaction*, *28*(3), 171–183.

Jalilifar, A., White, P., & Malekizadeh, N. (2017). 'Exploring Nominalization in Scientific Textbooks: A Cross-Disciplinary Study of Hard and Soft Sciences'. *International Journal of English Studies*, *17*(2), 1–20. https://doi.org/10.6018/ijes/2017/2/272781

Jalongo, M. R., Boyer, W., & Ebbeck, M. (2014). 'Writing for Scholarly Publication as "Tacit Knowledge": A Qualitative Focus Group Study of Doctoral Students in Education.' *Early Childhood Education Journal*, *42*(4), 241–250. https://doi.org/10.1007/s10643-013-0624-3

Jiang, F., & Hyland, K. (2015). '"The fact that": Stance nouns in disciplinary writing'. *Discourse Studies*, *17*(5), 529–550. https://doi.org/10.1177/1461445615590719

Johns, A. M. (2003). 'Genre and ESL/EFL composition instruction', in *Exploring the Dynamics of Second Language Writing,* Kroll, B. (ed.), (pp. 195–217). Cambridge: Cambridge University Press.

Johns, A. M. (2013). 'The History of English for Specific Purposes Research', in *The handbook of English for specific purposes*, Paltridge, B. & Starfield, S. (eds.), (pp. 5-30). Malden, Ma: John Wiley & Sons Inc. Available at: http://www.Nottingham.eblib.com/patron/FullRecord.aspx?p=1120617

Johnson, K., & Johnson, H. (2001). *Encyclopedic dictionary of applied linguistics: a handbook for language teaching.* Beijing: Foreign Language Press.

Jordan, R. R. (1997). *English for academic purposes: a guide and resource book for teachers.* Cambridge, UK; New York, NY: Cambridge University Press.

Kahn, J., & Wakin, D. J. (2007). 'Classical Music Looks Toward China With Hope'. *The New York Times*. Available at: http://www.nytimes.com/2007/04/03/arts/music/03class1.html

Kania, A. (2017). 'The Philosophy of Music' in *The Stanford Encyclopedia of Philosophy,* Zalta, E. N. (ed.). Stanford, CA:

Metaphysics Research Lab, Stanford University. Available at: https://plato.stanford.edu/archives/fall2017/entries/music/

Kaplan, R. B. (2010). 'Whence Applied Linguistics: The Twentieth century', in *The Oxford handbook of applied linguistics* (2nd ed.), Kaplan, R. B. (ed.), (pp. 3–33). Oxford: Oxford University Press.

Kashiha, H., & Heng, C. S. (2014). 'Discourse functions of formulaic sequences in academic speech across two disciplines'. *GEMA Online Journal of Language Studies*, *14*(2), 15–27.

Kellogg, R., & Raulerson, B. (2007). 'Improving the writing skills of college students.' *Psychonomic Bulletin & Review*, *14*(2), 237–242. https://doi.org/10.3758/BF03194058

Kennedy, C. (1983). 'An ESP Approach to EFL/ESL Teacher Training'. *ESP Journal*, *2*(1), 73–85. https://doi.org/10.1016/0272-2380(83)90024-0

Kennedy, G. D. (1998). *An introduction to corpus linguistics.* London: Longman.

Kivy, P. (1983). 'Platonism in Music: A Kind of Defense'. *Grazer Philosophische Studien*, *19*(1), 109–129.

Kivy, P. (1987). 'Platonism in Music: Another Kind of Defense'. *American Philosophical Quarterly*, *24*(3), 245–252.

Koester, A. (2010). 'Building small specialised corpora', in *The Routledge handbook of corpus linguistics,* O'Keeffe, A. & McCarthy, M. (eds.), (pp. 66–79). London: Routledge.

Kramer, L. (2003). 'Subjectivity Rampant! Music, Hermeneutics, and History', in *The cultural study of music: a critical introduction,* Clayton, M., Herbert, T., & Middleton, R. (eds.), (pp. 124–135). New York: Routledge.

Kramer, L. (2012). 'Oracular Musicology; or, Faking the Ineffable'. *Archiv Für Musikwissenschaft*, *69*(2), 101–109.

Lansdowne MS. 763 (c 1200- c1700). [manuscript] The British Library, Landsdowne Manuscripts. London.

Lansdowne, R. (2016). *Understanding music: philosophy and interpretation*. London: Bloomsbury Academic.

Law, W.-W., & Ho, W.-C. (2011). 'Music Education in China: In Search of Social Harmony and Chinese Nationalism'. *British Journal of Music Education*, *28*(3), 371–388. https://doi.org/10.1017/S0265051711000258

Lee, J. J., & Subtirelu, N. C. (2015). 'Metadiscourse in the classroom: A comparative analysis of EAP lessons and university lectures'. *English for Specific Purposes*, *37*, 52–62. https://doi.org/10.1016/j.esp.2014.06.005

Léon, J. (2006). 'Firth, John Rupert (1890-1960)', in *Encyclopedia of language & linguistics* (2nd ed.), [electronic resource], Brown, E. K. (ed.), (p. 484). Amsterdam: Elsevier. Available at: http://www.sciencedirect.com/science/referenceworks/9780080448541

Li, Jie, & Schmitt, N. (2009). 'The acquisition of lexical phrases in academic writing: A longitudinal case study'. *Journal of Second Language Writing*, *18*(2), 85–102. https://doi.org/10.1016/j.jslw.2009.02.001

Lillis, T. M., McKinney, C. (2003). *Analysing language in context: a student workbook.* Stoke-on-Trent: Trentham Books in association with The Open University.

Liu, D., & Nelson, R. (2016). 'Teaching language as a system', in *The Routledge Handbook of English language teaching,* Hall, G. (ed.), (pp. 413–427). Milton Park, Abingdon, Oxon; New York, NY: Routledge.

Llurda. (2016). '"Native speakers", English and ELT: Changing perspectives', in *The Routledge Handbook of English language teaching,* Hall, G. (ed.), (pp. 51–63). Milton Park, Abingdon, Oxon; New York, NY: Routledge.

Losada, C.C., 2014. 'Complex Multiplication, Structure, and Process: Harmony and Form in Boulez's Structures II'. *Music Theory Spectrum* 36, 86–120. https://doi.org/10.1093/mts/mtu005

Lovejoy, K. B. (1991). 'Cohesion and information strategies in academic writing: Analysis of passages in three disciplines'. *Linguistics and Education*, *3*(4), 315–343. https://doi.org/ 10.1016/0898-5898(91)90013-9

Lovejoy, K. B., & Lance, D. M. (1991). 'Information management and cohesion in the study of written discourse'. *Linguistics and Education*, *3*(3), 251–273. https://doi.org/10.1016/0898-5898(91)90010-G

Lu, W.-T. (2013). 'Confucius or Mozart? Community Cultural Wealth and Upward Mobility Among Children of Chinese Immigrants'. *Qualitative Sociology*, *36*(3), 303–321. https://doi.org/10.1007/s11133-013-9251-y

Malinowski, B. (1923). 'Supplement I', in *The Meaning of Meaning,* Ogden, C.K. & Richards, I.A. (pp. 296-336). New York: Harcourt, Brace and World, Inc.

Martin, J. R. (2012/2009). 'Boomer Dreaming: The Texture of Re-colonisation in a Lifestyle Magazine', in *Text Analysis,* Wang, Z. (ed.), (pp. 329–360). Shanghai: Shanghai Jiao Tong University Press.

Matheson, C., & Caplan, B. (2011). 'Ontology', in *The Routledge companion to philosophy and music* [electronic resource], Kania, A. & Gracyk, T. (eds.), (pp. 38–47). Abingdon, Oxon:

Routledge. Available at: http://www.myilibrary.com?id=304031

Matthiessen, C.M., 2015. Register in the round: registerial cartography. Functional Linguistics 2, 9. https://doi.org/10.1186/s40554-015-0015-8

Mauranen, A. (2003). 'The corpus of English as lingua franca in academic settings'. *Tesol Quarterly*, *37*(3), 513–527.

Mauranen, A., Hynninen, N., & Ranta, E. (2010). 'English as an academic lingua franca: The ELFA project'. *English for Specific Purposes*, *29*(3), 183–190. https://doi.org/10.1016/j.esp.2009.10.001

McEnery, T. (2004). 'Corpus Linguistics', in *The Oxford handbook of computational linguistics,* Mitkov, R. (ed.)*.* Oxford: Oxford University Press.

McEnery, T., & Hardie, A. (2012). *Corpus linguistics: method, theory and practice.* Cambridge: Cambridge University Press.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: an advanced resource book*. Abingdon: Routledge.

Melvin, S., & Cai, J. (2004). *Rhapsody in Red: How Western Classical Music Became Chinese*. New York: Algora Publishing.

Meunier, F. (2002). 'The pedagogical value of native and learner corpora in EFL grammar teaching', in *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching,* Granger, S., Hung, J., & Petch-Tyson, S. (eds.), (pp. 119–141). Amsterdam: John Benjamins Publishing Company.

Ministry of Education of the People's Republic of China/中华人民共和国教育部 （2016）.《中国留学回国就业蓝皮书2015》情况介绍. Retrieved March 16, 2018, from http://www.moe.edu.cn/jyb_xwfb/xw_fbh/moe_2069/xwfbh_2016n/xwfb_160325_01/160325_sfcl01/201603/t20160325_235214.html

Mollet, E., Wray, A., & Fitzpatrick, T. (2011). 'Accessing second-order collocation through lexical co-occurrence networks', in *The phraseological view of language: a tribute to John Sinclair,*

Sinclair, J. M., Herbst, T., Faulhaber, S. & Uhrig, P. (eds.), (pp. 87–121). Berlin ; Boston: De Gruyter Mouton.

Montefiore, C. S. (2014). 'Why piano-mania grips China's children'. *BBC Culture*. Available at: http://www.bbc.com/culture/story/ 20131022-piano-mania-grips-china

Morgan, M. H. (2014). *Speech communities.* Cambridge: Cambridge University Press.

Morton, T. (2010). 'Using a genre-based approach to integrating content and language in CLIL: The example of secondary history', in *Language use and language learning in CLIL classrooms*, Dalton-Puffer, C., Nikula, T., & Smit, U. (eds.), (pp. 81-104). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Myles, F. (2012). 'Complexity, accuracy and fluency', in *Dimensions of L2 performance and proficiency: complexity, accuracy and fluency in SLA,* Housen, A., Kuiken, F. & Vedder, I. (eds.), (pp. 71–93). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Nattiez, J.-J. (1987). *Musicologie générale et sémiologie*. Paris: Christian Bourgois éditeur.

Nattiez, J.-J. (1990). *Music and discourse: toward a semiology of music,* Abbate, C. (trans.). Princeton, NJ, London: Princeton University Press.

Nekrasova, T. M. (2009). 'English L1 and L2 Speakers' Knowledge of Lexical Bundles'. *Language Learning: A Journal of Research in Language Studies*, *59*(3), 647–686. https://doi.org/10.1111/j.1467-9922.2009.00520.x

Nelson, M. (2010). 'Building a written corpus: What are the basics?', in *The Routledge handbook of corpus linguistics,* O'Keeffe, A. & McCarthy, M. (eds.), (pp. 53–65). London: Routledge.

Nesselhauf, N. (2003). 'The Use of Collocations by Advanced Learners of English and Some Implications for Teaching'. *Applied Linguistics*, *24*(2), 223–242. https://doi.org/10.1093/applin/24.2.223

Nesselhauf, N. (2004). 'Learner corpora and their potential for language teaching', in *How to use corpora in language*

teaching, Sinclair, J. M. (ed.). Philadelphia, PA: John Benjamins Publishing Company. Available at: http://site.ebrary.com/lib/unnc/Doc?id=10056604

North, S. (2005). 'Disciplinary Variation in the Use of Theme in Undergraduate Essays'. *Applied Linguistics*, *26*(3), 431–452. https://doi.org/10.1093/applin/ami023

O'Donnell, M. B., Römer, U., & Ellis, N. C. (2015). 'The development of formulaic sequences in first and second language writing', in *Current issues in phraseology,* Hoffmann, S., Fischer-Starcke, B., & Sand, A. (eds.), (pp. 83–108). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Ortega, L. (2011). 'Second language acquisition', in *The Routledge handbook of applied linguistics* (1st ed.), Simpson, J. (ed.), (pp. 171–184). Milton Park, Abingdon; New York: Routledge.

Paltridge, B. (2013). 'Genre and English for Specific Purposes', in *The handbook of English for specific purposes,* Paltridge, B., & Starfield, S. (eds.), (pp. 347-366). Chichester, West Sussex: Wiley Blackwell.

Parkinson, J. (2013). 'Adopting academic values: Student use of that-complement clauses in academic writing'. *System*, *41*(2), 428–442. https://doi.org/10.1016/j.system.2013.03.002

Paul, T. & Diani, G. (2015). 'Introduction', in *English for academic purposes: approaches and implications,* Paul, T. & Diani, G. (eds.), (pp.1-8). Available at: https://ebookcentral.proquest.com/lib/nottingham/detail.action?docID=4534871

Paulus PP. VI. (1965). *Mysterium Fidei*. Libreria Editrice Vaticana. Available at: http://w2.vatican.va/content/paul-vi/la/encyclicals/documents/hf_p-vi_enc_03091965_mysterium.html

Pavlovová, M. (2013). *Complex linguistic analysis of musical discourse: The genre of concert notice*. Saarbrücken: LAP Lambert Academic Publishing.

Peregrin, J. (2012). 'The normative dimension of discourse', in *The Cambridge handbook of pragmatics,* Allan, K. & Jaszczolt, K. M. (eds.)*.* Cambridge: Cambridge University Press.

Pérez-Llantada, C. (2014). 'Formulaic language in L1 and L2 expert academic writing: Convergent and divergent usage'. *Journal of English for Academic Purposes*, *14*, 84–94.

Pérez-Sobrino, P., & Julich, N. (2014). 'Let's Talk Music: A Corpus-Based Account of Musical Motion'. *Metaphor and Symbol*, *29*(4), 298–315. https://doi.org/10.1080/10926488.2014.948800

Peters, E., & Pauwels, P. (2015). 'Learning academic formulaic sequences'. *Journal of English for Academic Purposes*, *20*, 28–39. https://doi.org/10.1016/j.jeap.2015.04.002

Phillipson, R. (1992). *Linguistic imperialism.* Oxford: Oxford University Press.

Poultney, D. (1995). *Studying Music History: Learning, Reasoning, and Writing About Music History and Literature* (2nd ed.). London: Pearson.

Povolná, R. (2012). 'Causal and contrastive discourse markers in novice academic writing'. *Brno Studies in English*, *38*(2), 131–148. https://doi.org/10.5817/BSE2012-2-8

Pratt, D. (2011). *Modelling written communication: a new systems approach to modelling in the social sciences.* Dordrecht; New York: Springer Verlag.

Qin, J. (2014). 'Use of Formulaic Bundles by Non-Native English Graduate Writers and Published Authors in Applied Linguistics'. *System: An International Journal of Educational Technology and Applied Linguistics*, *42*, 220–231.

Reppen, R. (2010). 'Building a corpus: What are the key considerations', in *The Routledge handbook of corpus linguistics,* O'Keeffe, A. & McCarthy, M. (eds.), (pp. 31–37). London: Routledge.

Rohrbaugh, G. (2003). 'Artworks as Historical Individuals'. *European Journal of Philosophy*, *11*(2), 177–205. https://doi.org/10.1111/1468-0378.00182

Sanders, T., & Sanders, J. (2006). 'Text and Text Analysis', in the *Encyclopedia of language & linguistics* [electronic resource], Brown, E. K. (ed.), (2nd ed., pp. 597–607). Amsterdam: Elsevier. Available at: https://doi-

org.ezproxy.nottingham.edu.cn/10.1016/B0-08-044854-2/00557-5

Savage, P. E., Brown, S., Sakai, E., & Currie, T. E. (2015). 'Statistical universals reveal the structures and functions of human music'. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(29), 8987–8992. https://doi.org/10.1073/pnas.1414495112

Schmitt, N., Grandage, S., & Adolphs, S. (2004). 'Are corpus-derived recurrent clusters psycholinguistically valid?', in *Formulaic sequences: acquisition, processing, and use* [electronic resource], Schmitt, N. (ed.), (pp. 127–152). Amsterdam; Philadelphia: John Benjamins Publishing Company'. Available at: https://ebookcentral.proquest.com/lib/nottingham/detail.action?docID=622575

SCImago, n.d. SCImago, (n.d.). SJR — SCImago Journal & Country Rank [Portal]. [WWW Document]. SJR Scimago Journal & Country Rank. URL http://www.scimagojr.com (accessed 9.30.19).

Scott, M. (2017). WordSmith Tools (Version 7). Stroud: Lexical Analysis Software. Available at: http://lexically.net/wordsmith/

Sharpe, R. A. (2004). *Philosophy of music: an introduction,* [electronic resource]. Stocksfield UK: Acumen. Available at: https://ebookcentral.proquest.com/lib/nottingham/detail.action?docID=1968917

Simpson, J. (2011). 'Introduction: Applied linguistics in the contemporary world', in *The Routledge handbook of applied linguistics* [electronic resource], Simpson, J. (ed.), (pp. 1–7). Milton Park, Abingdon: Routledge. Available at: http://Nottingham.eblib.com/patron/FullRecord.aspx?p=574657

Simpson-Vlach, R., & Ellis, N. C. (2010). 'An Academic Formulas List: New Methods in Phraseology Research'. *Applied Linguistics*, *31*(4), 487–512. https://doi.org/10.1093/applin/amp058

Sinclair, J. M. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Sinclair, J. M. (2004a). *English collocation studies: the OSTI report*. London; New York: Continuum.

Sinclair, J. M. (2004b). *Trust the text: language, corpus and discourse*. London; New York, N.Y.: Routledge.

Sinclair, J. M. (2005). 'Corpus and Text-Basic Principles', in *Developing linguistic corpora: a guide to good practice,* Wynne, M. (ed.), (pp. 1–16). Oxford: Oxbow Books on behalf of the Arts and Humanities Data Service; Oakville, CT: Available direct in the US from David Brown Book Company.

Skehan, P. (2016). 'Sidesteps towards applied linguistics', in *Becoming and Being an Applied Linguist,* Ellis, R. (ed.), (pp. 89–117). Amsterdam: John Benjamins Publishing Company.

Slabakova, R. (2014). 'L2 Semantics', in *The Routledge handbook of second language acquisition,* Gass, S. M. & Mackey, A. (eds.), (pp. 127–146). London: Routledge, Taylor & Francis Group.

Slabakova, R. (2016). *Second language acquisition.* Oxford: Oxford University Press.

Smith, N. (2017). 'Who Has the World's No. 1 Economy? Not the U.S.' Bloomberg View. Available at: https://www.bloomberg.com/view/articles/2017-10-18/who-has-the-world-s-no-1-economy-not-the-u-s

Society for Music Theory, 2018. SMT: The Society for Music Theory [WWW Document]. URL https://societymusictheory.org/music-theory-spectrum (accessed 9.30.19).

Solly, M. (2016). *The stylistics of professional discourse*. Edinburgh: Edinburgh University Press.

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). 'Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section'. *Journal of English for Academic Purposes*, *12*(3), 214–225. https://doi.org/10.1016/j.jeap.2013.05.002

Starfield, S. (2016). 'English for specific purposes', in *The Routledge Handbook of English language teaching,* Hall, G. (ed.), (pp. 150–163). Milton Park, Abingdon, Oxon; New York, NY: Routledge.

Stubbs, M. (2007a). 'An example of frequent English phraseology: Distribution, structures and functions', in *Corpus linguistics 25 years on* [electronic resource], Facchinetti, R. (ed.), (pp. 89–105). Amsterdam; New York: Rodopi. Available at: https://ebookcentral.proquest.com/lib/nottingham/detail.action?docID=556853

Stubbs, M. (2007c). 'Quantitative data on multi-word sequences in English: the case of the word world', in *Text, discourse and corpora: theory and analysis; with an introduction by John Sinclair,* (pp. 163–190). London; New York: Continuum.

Suzuki, D.T. (2002). *Mysticism: Christian and Buddhist* (2nd ed.). London, New York: Routledge.

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge: Cambridge University Press.

Tannen, D., Hamilton, H. E., & Schiffrin, D. (2015). *The handbook of discourse analysis* (2nd ed.). Malden, MA: John Wiley & Sons, Inc.

Tarone, E., Dwyer, S., Gillette, S., & Icke, V. (1981). 'On the use of the passive in two astrophysics journal papers'. *The ESP Journal, 1*(2), 123–140. https://doi.org/10.1016/0272-2380(81)90004-4

Teubert, W. (2004). 'Language and corpus linguistics', in *Lexicology and corpus linguistics: an introduction,* Halliday, M. A. K., et al. (eds.), (pp. 73–112). London: Continuum.

Teubert, W., & Čermáková, A. (2004). 'Directions in corpus linguistics', in Lexicology and corpus linguistics: an introduction, Halliday, M. A. K., et al. (eds.), (pp. 113–165). London: Continuum.

The World Bank. (2018). 'Data: China'. Retrieved March 16, 2018, from https://data.worldbank.org/country/china

Thomas, J. (2015). 'Stealing a march on collocation: Deriving extended collocations from full text for student analysis and synthesis', in *Multiple affordances of language corpora for data-driven learning,* Lenko-Szymanska, A., & Boulton, A. (eds.), (pp. 85–108). Amsterdam; Philadelphia: John Benjamins Publishing Company.

Thompson, G. (2014). *Introducing functional grammar* [electronic resource], (3rd ed.). Abingdon, Oxon: Routledge. Available at: http://Nottingham.eblib.com/patron/FullRecord.aspx?p=1319003

Thompson, G., & Hunston, S. (2006). 'System and corpus: two traditions with a common ground', in *System and corpus: exploring connections,* Thompson, G., & Hunston, S. (eds.), (pp. 1–14). London: Equinox.

Thornbury, S. (2010). 'What can a corpus tell us about discourse?', in *The Routledge handbook of corpus linguistics,* O'Keeffe, A., & McCarthy, M. (eds.), (pp. 270–287). London: Routledge.

Titscher, S., Meyer, M., Wodak, R., & Vetter, E. (2000). Methods of text and discourse analysis, Jenner, B. (trans.). London: SAGE Publications.

Todd, R. W. (2016). *Discourse topics.* Amsterdam; Philadelphia: Based on the author's PhD thesis entitled "Topics in classroom discourse", University of Liverpool, Liverpool UK, 2003.

Tuan, Y., (2009). *Religion: From lace to Placelessness* [text by Yi-Fu Tuan]. 1st ed. Chicago, Ill.: Center for American Places at Columbia College Chicago: Distributed by the University of Chicago Press, Chicago, Ill.

Vickers, C. H. (2010). 'Language competence and the construction of expert–novice in NS–NNS interaction'. *Journal of Pragmatics*, *42*(1), 116–138. https://doi.org/10.1016/j.pragma.2009.05.010

Vyatkina, N., Hirschmann, H., Golcher, F., 2015. Syntactic modification at early stages of L2 German writing development: A longitudinal learner corpus study. Journal of Second Language Writing 29, 28–50. https://doi.org/10.1016/j.jslw.2015.06.006

Widdowson, H. G. (1983). *Learning Purpose and Language Use*. Oxford: Oxford University Press.

Wiley Online Library: Music Analysis [WWW Document], n.d. Browse Issue Archive. URL https://onlinelibrary.wiley.com/loi/14682249/year/2018 (accessed 9.30.19).

Winberg, C., Van Der Geest, T., Lehman, B., & Nduna, J. (2010). 'Teaching technical writing in multilingual contexts: A meta-analysis'. *Southern African Linguistics and Applied Language Studies*, *28*(3), 299–308. https://doi.org/10.2989/16073614.2010.545032

Wingell, R. J. (2008). *Writing about Music: An Introductory Guide* (4th ed.). London: Pearson.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. London: Kegan Paul.

Wood, D. (2015). *Fundamentals of formulaic language: an introduction.* London: Bloomsbury.

Wray, A. (1999). 'Formulaic Language in Learners and Native Speakers'. *Language Teaching: The International Abstracting Journal for Language Teachers, Educators and Researchers*, *32*(4), 213–231.

Wray, A. (2008). *Formulaic language: pushing the boundaries.* Oxford: Oxford University Press.

Wray, A. (2012). 'What Do We (Think We) Know About Formulaic Language? An Evaluation of the Current State of Play'. *Annual Review of Applied Linguistics*, *32*, 231–254. https://doi.org/10.1017/S026719051200013X

Wray, A., & Hunston. (2000). 'The Functions of Formulaic Language: An Integrated Model'. *Language & Communication*, *20*(1), 1–28.

Yasuda, S. (2011). 'Genre-Based Tasks in Foreign Language Writing: Developing Writers' Genre Awareness, Linguistic Knowledge, and Writing Competence'. *Journal of Second Language Writing*, *20*(2), 111–133. https://doi.org/10.1016/j.jslw.2011.03.001

Young, J. (2011). 'The ontology of musical works: A philosophical pseudo-problem'. *Frontiers of Philosophy in China*, *6*(2), 284–297. https://doi.org/10.1007/s11466-011-0139-1

Yuldashev, A., Fernandez, J., & Thorne, S. L. (2013). 'Second Language Learners' Contiguous and Discontiguous Multi-Word Unit Use Over Time'. *The Modern Language Journal*, *97*(S1), 31–45. https://doi.org/10.1111/j.1540-4781.2012.01420.x

Zbikowski, L. M. (1998). 'Metaphor and Music Theory: Reflections from Cognitive Science'. *Music Theory Online*, *4*(1), 1–11.

Zbikowski, L. M. (2008). 'Metaphor and Music', in *The Cambridge handbook of metaphor and thought,* Gibbs, R. W. (ed.), (pp. 502–524). Cambridge: Cambridge University Press.

References:

Expert Corpus Articles

Adams, K. (2015). 'What Did Danger Mouse Do? The Grey Album and Musical Composition in Configurable Culture.' *Music Theory Spectrum* 37, 7–24. https://doi.org/10.1093/mts/mtv004

Alexander, P. (2018). '"Our city of love and of slaughter": Berlin klezmer and the politics of place.' *Ethnomusicology Forum* 27, 25–47. https://doi.org/10.1080/17411912.2018.1460616

Allis, M. (2014). 'Bantock and Southey: Musical Otherness and Fatalism in Thalaba the Destroyer.' *Music and Letters* 95, 39–69. https://doi.org/10.1093/ml/gct132

Argentino, J. (2018). 'Schoenberg and Liszt: Hexatonic Collections across the Great Tonal Divide.' *Music Analysis* 37, 184–202. https://doi.org/10.1111/musa.12093

Atkinson, P. (2016). '"An Ireland Built Anew": Bax's "Tintagel" and the Easter Rising.' *Music and Letters* 97, 100–135. https://doi.org/10.1093/ml/gcw013

Bain, J. (2015). 'Was Hildegard Forgotten?' *Journal of Musicological Research* 34, 1–30. https://doi.org/10.1080/01411896.2015.985520

Ballantine, C. (2017). 'Song, Memory, Power, and the South African Archive.' *The Musical Quarterly* 99, 60–80. https://doi.org/10.1093/musqtl/gdw011

Beng, T. S. (2016). 'Cosmopolitan Identities: Evolving Musical Cultures of the Straits-born Chinese of Pre-World War II Malaya.' *Ethnomusicology Forum* 25, 35–57. https://doi.org/10.1080/17411912.2016.1151813

Beng, T. S., & Rao, N. Y. (2016). 'Introduction—emergent Sino-soundscapes: Musical Pasts, Transnationalism and Multiple Identities.' *Ethnomusicology Forum* 25, 4–13. https://doi.org/10.1080/17411912.2016.1151812

Benoit-Otis, M.-H., & Quesney, C. (2017). 'A Nazi Pilgrimage to Vienna? The French Delegation at the 1941 "Mozart Week of the German Reich."' *The Musical Quarterly* 99, 6–59. https://doi.org/10.1093/musqtl/gdw010

Bevilacqua, G. (2016). 'The Earliest Source of Notre-Dame Polyphony? A New Conductus Fragment from the Early Thirteenth Century.' *Music and Letters* 97, 1–41. https://doi.org/10.1093/ml/gcw008

Blake, D. K. (2016). 'University Geographies and Folk Music Landscapes.' *The Journal of Musicology* 33, 92. https://doi.org/10.1525/jm.2016.33.1.92

Bohlman, A. F., & Scheding, F. (2015). 'Hanns Eisler on the Move: Tracing Mobility in the "Reisesonate."' *Music and Letters* 96, 77–98. https://doi.org/10.1093/ml/gcu122

Boss, J. (2018). 'Schoenberg's Opus 33B and the Problem of its Contrasting "Continuation" and Second Theme.' *Music Analysis* 37, 203–242. https://doi.org/10.1111/musa.12105

Bottge, K. M. (2015). 'Lessons in "Pure Visibility" (reine Sichtbarkeit): Victor Hammer's Correspondence with Heinrich Schenker.' *Music Theory Spectrum* 37, 25–50. https://doi.org/10.1093/mts/mtu019

Bradley, C. A. (2014). 'Comparing Compositional Process in Two Thirteenth-Century Motets: Pre-existent Materials in Deus omnium/REGNAT and Ne m'oubliez mie/DOMINO.' *Music Analysis* 33, 263–290. https://doi.org/10.1111/musa.12029

Brand, B. (2017). 'Plainsong New and Old: The Versified Office for St Ansanus of Siena.' *Music and Letters* 98, 1–31. https://doi.org/10.1093/ml/gcx001

Branda, E. (2017). 'Speaking German, Hearing Czech, Claiming Dvořák.' *Journal of the Royal Musical Association* 142, 109–136. https://doi.org/10.1080/02690403.2017.1286129

Breyley, G. J. (2016). 'Between the Cracks: Street Music in Iran.' *Journal of Musicological Research* 35, 72–81. https://doi.org/10.1080/01411896.2016.1165051

Brown, A. (2014). 'Invented Plainsongs in Keyboard Settings Ascribed to Tomkins and Gibbons.' *Music and Letters* 95, 1–16. https://doi.org/10.1093/ml/gcu009

Browning, J. (2016). 'Assembled Landscapes.' *The Journal of Musicology* 33, 70. https://doi.org/10.1525/jm.2016.33.1.70

Buch, E. (2017). 'Mahler's Fifth, Daniel Barenboim, and the Argentine Dictatorship: On Music, Meaning, and Politics.' *The Musical Quarterly* 100, 122–154. https://doi.org/10.1093/musqtl/gdx015

Butler, K. (2016). 'Changing Attitudes Towards Classical Mythology and their Impact on Notions of the Powers of Music in Early Modern England.' *Music and Letters* 97, 42–60. https://doi.org/10.1093/ml/gcw015

Butt, J. (2018). 'Monteverdi, the 1610 Vespers and the Beginnings of the Modern Musical Work.' *Journal of the Royal Musical Association* 143, 21–50. https://doi.org/10.1080/02690403.2018.1434328

Caplin, W. E. (2018). 'Beyond the Classical Cadence: Thematic Closure in Early Romantic Music.' *Music Theory Spectrum* 40, 1–26. https://doi.org/10.1093/mts/mty002

Caplin, W. E., & Martin, N. J. (2016). 'The "Continuous Exposition" and the Concept of Subordinate Theme.' *Music Analysis* 35, 4–43. https://doi.org/10.1111/musa.12060

Carter-Ényì, A. (2018). '"Lùlù fún wọn": Oríkì in Contemporary Culture.' *Ethnomusicology* 62, 83–103. https://doi.org/10.5406/ethnomusicology.62.1.0083

Ceballos, S. G. (2017). 'Sympathizing with C. P. E. Bachs Empfindungen.' *The Journal of Musicology* 34, 1. https://doi.org/10.1525/jm.2017.34.01.1

Chew, G. (2018). 'The Enchanting Music of Blades And Bullets: Visions of Violence in Janáček's World-War Symphonic Poems.' *Music and Letters* 99, 16–44. https://doi.org/10.1093/ml/gcy011

Clarke, E., Doffman, M., & Timmers, R. (2016). 'Creativity, Collaboration and Development in Jeremy Thurlow's Ouija for Peter Sheppard Skærved.' *Journal of the Royal Musical Association* 141, 113–165. https://doi.org/10.1080/02690403.2016.1151240

Cohen, B. (2014). 'Limits of National History: Yoko Ono, Stefan Wolpe, and Dilemmas of Cosmopolitanism.' *The Musical Quarterly* 97, 181–237. https://doi.org/10.1093/musqtl/gdu008

Cook, J. (2015). 'The Style of Walter Frye and an Anonymous Mass in the Lucca Choirbook.' *Music and Letters* 96, 1–27. https://doi.org/10.1093/ml/gcv004

Cook, N. (2017). 'Inventing Tradition: Webern's Piano Variations in Early Recordings.' *Music Analysis* 36, 163–215. https://doi.org/10.1111/musa.12094

Cormac, J. (2017). 'From Satirical Piece to Commercial Product: The Mid-Victorian Opera Burlesque and its Bourgeois Audience.' *Journal of the Royal Musical Association* 142, 69–108. https://doi.org/10.1080/02690403.2017.1286124

Čurda, M. (2016). '"From the Monkey Mountains": The Body, the Grotesque and Carnival in the Music of Pavel Haas.' *Journal of the Royal Musical Association* 141, 61–112. https://doi.org/10.1080/02690403.2016.1151235

Cypess, R. (2017). '"It Would Be without Error": Automated Technology and the Pursuit of Correct Performance in the French Enlightenment.' *Journal of the Royal Musical Association* 142, 1–29. https://doi.org/10.1080/02690403.2017.1286115

Dave, N. (2014). 'The Politics of Silence: Music, Violence and Protest in Guinea.' *Ethnomusicology* 58, 1–29. https://doi.org/10.5406/ethnomusicology.58.1.0001

Deeming, H. (2014). 'Music and Contemplation in the Twelfth-Century Dulcis Jesu memoria.' *Journal of the Royal Musical Association* 139, 1–39. https://doi.org/10.1080/02690403.2014.886410

DeFord, R.I. (2016). 'Diminution in the Theory of Johannes de Muris and His Followers.' *Music Theory Spectrum* 38, 58–84. https://doi.org/10.1093/mts/mtw002

Diaz, J. D. (2017). 'Between Repetition and Variation: A Musical Performance of Malícia in Capoeira.' *Ethnomusicology Forum* 26, 46–68. https://doi.org/10.1080/17411912.2017.1309297

Diettrich. B. (2018). '"Summoning Breadfruit" and "Opening Seas": Toward a Performative Ecology in Oceania.' *Ethnomusicology* 62, 1–27. https://doi.org/10.5406/ethnomusicology.62.1.0001

Duane, B. (2017). 'Thematic and Non-Thematic Textures in Schubert's Three-Key Expositions.' *Music Theory Spectrum* 39, 36–65. https://doi.org/10.1093/mts/mtx007

Frankenbach, C. (2017). 'Dancing to Beethoven in Wilhelmine Germany.' *The Journal of Musicology* 34, 71. https://doi.org/10.1525/jm.2017.34.01.71

Frolova-Walker, M. (2018). 'An Inclusive History for a Divided World?' *Journal of the Royal Musical Association* 143, 1–20. https://doi.org/10.1080/02690403.2018.1434308

Gawboy, A. (2017). '"Oneness Through the Sharpening of Contradictions": Theosophical Polarity in Scriabin's Late Harmonic Practice.' *Journal of Musicological Research* 36, 181–207. https://doi.org/10.1080/01411896.2017.1341270

Gidal, M. M. (2014). 'Musical Boundary-Work: Ethnomusicology, Symbolic Boundary Studies, and Music in the Afro-Gaucho

Religious Community of Southern Brazil.' *Ethnomusicology* 58, 83–109. https://doi.org/10.5406/ethnomusicology.58.1.0083

Glatthorn, A. (2018). 'In the Name of the Emperor.' *The Journal of Musicology* 35, 1. https://doi.org/10.1525/jm.2018.35.1.1

Goldmark, D. (2015). '"Making Songs Pay": Tin Pan Alley's Formula for Success.' *The Musical Quarterly* 98, 3–28. https://doi.org/10.1093/musqtl/gdv012

Graves, E. (2017). 'The Marketplace of Devotional Song: Cultural Economies of Exchange in Bengali Padāvalī-Kīrtan.' *Ethnomusicology* 61, 52–86. https://doi.org/10.5406/ethnomusicology.61.1.0052

Greenberg, Y. (2018). '"Ordo ab Chao": The Fugue as Chaos in the Early Twentieth century.' *Music and Letters* 99, 74–103. https://doi.org/10.1093/ml/gcy010

Greenspan, C. (2015). 'Richard Rodgers: Collaborator.' *The Musical Quarterly* 98, 81–99. https://doi.org/10.1093/musqtl/gdv008

Gregor, N. (2015). 'Music, Memory, Emotion: Richard Strauss and The Legacies of War.' *Music and Letters* 96, 55–76. https://doi.org/10.1093/ml/gcu110

Grimley, D. M. (2016). 'Music, Landscape, and the Sound of Place.' *The Journal of Musicology* 33, 11. https://doi.org/10.1525/jm.2016.33.1.11

Guthrie, K. (2014). 'Propaganda Music in Second World War Britain: John Ireland's Epic March.' *Journal of the Royal Musical Association* 139, 137–175. https://doi.org/10.1080/02690403.2014.886430

Hand, R. (2018). 'Schools and Families as Institutions of Learning in Central Javanese gamelan.' *Ethnomusicology Forum* 27, 68–87. https://doi.org/10.1080/17411912.2018.1463550

Helgert, L. E. (2014). 'Lukas Foss's American Works as Expressions of the Immigrant Experience.' *Journal of Musicological Research* 33, 315–351. https://doi.org/10.1080/01411896.2014.954511

Hepokoski, J. A. (2016). 'Sonata Theory, Secondary Themes and Continuous Expositions: Dialogues with Form-Functional Theory.' *Music Analysis* 35, 44–74. https://doi.org/10.1111/musa.12061

Hofman, A. (2015). 'Music (as) Labour: Professional Musicianship, Affective Labour and Gender in Socialist Yugoslavia.' *Ethnomusicology Forum* 24, 28–50. https://doi.org/10.1080/17411912.2015.1009479

Horton, J. (2018). 'Form and Orbital Tonality in the Finale of Bruckner's Seventh Symphony.' *Music Analysis* 37, 271–309. https://doi.org/10.1111/musa.12124

Howland, P. (2015). 'Formal Structures in Post-Tonal Music.' *Music Theory Spectrum* 37, 71–97. https://doi.org/10.1093/mts/mtv011

Huebner, S. (2017). 'Classical Wagnerism.' *The Journal of Musicology* 34, 115. https://doi.org/10.1525/jm.2017.34.01.115

Huebner, S. (2014). 'Ravel's Politics.' *The Musical Quarterly* 97, 66–97. https://doi.org/10.1093/musqtl/gdu003

Jarjour, T. (2015). 'Ḥasho: Music Modality and the Economy of Emotional Aesthetics.' *Ethnomusicology Forum* 24, 51–72. https://doi.org/10.1080/17411912.2015.1018918

Jerold, B. (2018). 'A Vindication of Ferdinand Hiller.' *Journal of Musicological Research* 37, 141–165. https://doi.org/10.1080/01411896.2018.1453694

Johansson, M. (2017). 'Paraphrase This: A Note on Improvisation.' *Ethnomusicology* Forum 26, 26–45. https://doi.org/10.1080/17411912.2017.1302810

Johansson, M., & Berge, O. K. (2014). 'Who Owns an Interpretation? Legal and Symbolic Ownership of Norwegian Folk Music.' *Ethnomusicology* 58, 30–53. https://doi.org/10.5406/ethnomusicology.58.1.0030

Johnston, B. (2014). 'Salome's Grotesque Climax and Its Implications.' *Music Theory Spectrum* 36, 34–57. https://doi.org/10.1093/mts/mtu004

Jung, E.-Y. (2014). 'Transnational Migrations and YouTube Sensations: Korean Americans, Popular Music, and Social Media.' *Ethnomusicology* 58, 54–82. https://doi.org/10.5406/ethnomusicology.58.1.0054

Kallis, V. (2015). 'Pitch Organisation in Scriabin's Tenth Sonata.' *Music Analysis* 34, 3–46. https://doi.org/10.1111/musa.12039

Kartomi, M. (2017). 'From Agricultural Tool to Identity Symbol: Musical and Socio-Political Change in the Pestle-and-Mortar Music of the Natuna Archipelago, Indonesia.' *Ethnomusicology* 61, 87–114. https://doi.org/10.5406/ethnomusicology.61.1.0087

Katz, M. (2014). 'Sites of Memory in Hindustani Music: Yusuf Ali Khan and the Sitar Shops of Lucknow.' *Ethnomusicology Forum* 23, 67–93. https://doi.org/10.1080/17411912.2013.864537

Kawabata, M. (2015). 'Playing the "Unplayable": Schoenberg, Heifetz, and the Violin Concerto, Op. 36.' *Journal of*

Musicological Research 34, 31–50. https://doi.org/10.1080/01411896.2014.956494

Keegan-Phipps, S. (2017). 'Identifying the English: Essentialism and Multiculturalism in Contemporary English Folk Music.' *Ethnomusicology Forum* 26, 3–25. https://doi.org/10.1080/17411912.2017.1302809

Keightley, K. (2015). 'Hogan's Tin Pan Alley: R. F. Outcault and Popular Sheet Music.' *The Musical Quarterly* 98, 29–56. https://doi.org/10.1093/musqtl/gdv011

Kieffer, A. (2016). 'Riemann in France: Jean Marnold and the "Modern" Music-Theoretical Ear.' *Music Theory Spectrum* 38, 1–15. https://doi.org/10.1093/mts/mtv038

Kim, M. (2016). 'The Amalekite Case and the Portrayal of Saul in Handel and Jennens's Oratorio Saul (1739).' *Journal of Musicological Research* 35, 45–60. https://doi.org/10.1080/01411896.2016.1116348

Kwon, D.L. (2015). '"Becoming One": Embodying Korean P'ungmul Percussion Band Music and Dance through Site-Specific

Intermodal Transmission.' *Ethnomusicology* 59, 31–60. https://doi.org/10.5406/ethnomusicology.59.1.0031

Lampert, V. (2018). 'Benny Goodman and Bartók's Contrasts: An Extraordinary Collaboration.' *The Musical Quarterly* 101, 6–34. https://doi.org/10.1093/musqtl/gdy010

Latour, M. (2015). 'Disciplining Song in Sixteenth-Century Geneva.' *The Journal of Musicology* 32, 1. https://doi.org/10.1525/jm.2015.32.1.1

Lawson, F. R. S. (2014). 'Is Music an Adaptation or a Technology? Ethnomusicological Perspectives from the Analysis of Chinese Shuochang.' *Ethnomusicology* Forum 23, 3–26. https://doi.org/10.1080/17411912.2013.875786

Leech-Wilkinson, D. (2015). 'Cortot's Berceuse.' *Music Analysis* 34, 335–363. https://doi.org/10.1111/musa.12054

Leistra-Jones, K. (2018). 'Hans von Bülow and the Confessionalization of Kunstreligion.' *The Journal of Musicology* 35, 42. https://doi.org/10.1525/jm.2018.35.1.42

Leistra-Jones, K. (2016). '"The Deeps have Music Soft and Low": Sounding the Ocean in Elgar's Sea Pictures.' *Music and Letters* 97, 61–99. https://doi.org/10.1093/ml/gcw009

Levy, D.S. (2014). 'Parsifal in Yiddish? Why Not?' *The Musical Quarterly* 97, 140–180. https://doi.org/10.1093/musqtl/gdu007

Lewis, H. (2018). 'Decasia as Audiovisual Elegy.' *Journal of Musicological Research* 37, 189–208. https://doi.org/10.1080/01411896.2018.1477399

Long, M. K. (2018). 'Cadential Syntax and Tonal Expectation in Late Sixteenth-Century Homophony.' *Music Theory Spectrum* 40, 52–83. https://doi.org/10.1093/mts/mty004

Losada, C. C. (2014). 'Complex Multiplication, Structure, and Process: Harmony and Form in Boulez's Structures II.' *Music Theory Spectrum* 36, 86–120. https://doi.org/10.1093/mts/mtu005

Loughridge, D. (2016). 'Making, Collecting and Reading Music Facsimiles before Photography.' *Journal of the Royal*

Musical Association 141, 27–59. https://doi.org/10.1080/02690403.2016.1151232

Lu, H.-C. T. (2016). 'Cultural Homogeneity, Embodied Empathy: Incorporating Musical Pasts among Burmese Chinese Peoples Worldwide.' *Ethnomusicology Forum* 25, 14–34. https://doi.org/10.1080/17411912.2016.1151811

MacGregor, E. (2018). 'Listening for the Intimsphäre: Recovering Berlin 1933 through Hans Pfitzner's Symphony in C-sharp Minor.' *The Musical Quarterly* 101, 35–75. https://doi.org/10.1093/musqtl/gdy008

Machin-Autenrieth, M. (2015). 'Flamenco ¿Algo Nuestro? (Something of Ours?): Music, Regionalism and Political Geography in Andalusia, Spain.' *Ethnomusicology Forum* 24, 4–27. https://doi.org/10.1080/17411912.2014.966852

Manuel, P. (2015). 'The Intermediate Sphere in North Indian Music Culture: Between and Beyond "Folk" and "Classical."' *Ethnomusicology* 59, 82–115. https://doi.org/10.5406/ethnomusicology.59.1.0082

Mark, C. (2018). 'Bridge and Britten, Britten and Bridge.' *Music and Letters* 99, 45–73. https://doi.org/10.1093/ml/gcy001

Martin, H. (2018). 'Prolongation and Its Limits: The Compositions of Wayne Shorter.' *Music Theory Spectrum* 40, 84–105. https://doi.org/10.1093/mts/mty006

Johansson, M. (2017). 'Non-Isochronous Musical Meters: Towards a Multidimensional Model.' *Ethnomusicology* 61, 31–51. https://doi.org/10.5406/ethnomusicology.61.1.0031

Matsue, J. M. (2016). 'Drumming to One's Own Beat: Japanese Taiko and the Challenge to Genre.' *Ethnomusicology* 60, 22–52. https://doi.org/10.5406/ethnomusicology.60.1.0022

Mayes, C. (2014). 'Eastern European National Music as Concept and Commodity at the Turn of the Nineteenth Century.' *Music and Letters* 95, 70–91. https://doi.org/10.1093/ml/gct130

McGuire, C. P. (2018). 'Unisonance in Kung Fu Film Music, or the Wong Fei-hung Theme Song as a Cantonese Transnational Anthem.' *Ethnomusicology Forum* 27, 48–67. https://doi.org/10.1080/17411912.2018.1463549

Meinhart, M. (2014). 'Memory and Private Mourning in an English Country House during the First World War: Lady Alda Hoare's Musical Shrine to a Lost Son.' *Journal of Musicological Research* 33, 39–95. https://doi.org/10.1080/01411896.2014.876489

Meneses, J. D. D. (2016). 'Listening with the Body: An Aesthetics of Spirit Possession Outside the Terreiro.' *Ethnomusicology* 60, 89–124. https://doi.org/10.5406/ethnomusicology.60.1.0089

Mercer-Taylor, P. (2015). 'Listening for the "Still Small Voice" of Mendelssohn's Domestic Elijah.' *The Journal of Musicology* 32, 40. https://doi.org/10.1525/jm.2015.32.1.40

Merli, P. (2018). 'La Scala in the Aftermath of the Liberation, 25 April 1945 to 22 June 1946.' *The Musical Quarterly* 100, 155–198. https://doi.org/10.1093/musqtl/gdy002

Meyers, J. P. (2015). 'Still Like That Old Time Rock and Roll: Tribute Bands and Historical Consciousness in Popular Music.' *Ethnomusicology* 59, 61–81. https://doi.org/10.5406/ethnomusicology.59.1.0061

Mills, S., & Park, S.-H. (2017). 'The Musical Faces of South Korea's East Coast Shaman Tradition: an Exploration into Personal Style Formation and Expression.' *Ethnomusicology Forum* 26, 69–92. https://doi.org/10.1080/17411912.2017.1305283

Mould, S. (2014). 'Fidelity to "Fidelio": The Convention of Unwritten Appoggiaturas in Beethoven's Only Opera.' *Music and Letters* 95, 17–38. https://doi.org/10.1093/ml/gcu008

Mount, A. (2015). 'Grasp the Weapon of Culture! Radical Avant-Gardes and the Los Angeles Free Press.' *The Journal of Musicology* 32, 115. https://doi.org/10.1525/jm.2015.32.1.115

Mugmon, M. (2018). 'An Imperfect Mahlerite.' *The Journal of Musicology* 35, 76. https://doi.org/10.1525/jm.2018.35.1.76

Mugmon, M. (2014). 'Patriotism, Art, and "The Star-Spangled Banner" in World War I: A New Look at the Karl Muck Episode.' *Journal of Musicological Research* 33, 4–26. https://doi.org/10.1080/01411896.2014.879028

Muller, S. (2016). 'Apartheid Aesthetics and Insignificant Art.' *The Journal of Musicology* 33, 45. https://doi.org/10.1525/jm.2016.33.1.45

Murphy, E. (2017). '"Liveridge is in Ireland": Richard Leveridge and the Earliest Surviving Dublin Birthday Odes.' *Music and Letters* 98, 32–73. https://doi.org/10.1093/ml/gcx012

Negley, B. (2017). 'The Ford Foundation Symphony Orchestra Program.' *Journal of Musicological Research* 36, 115–142. https://doi.org/10.1080/01411896.2017.1294959

Neimoyer, S. (2014). 'After the Rhapsody.' *The Journal of Musicology* 31, 91. https://doi.org/10.1525/jm.2014.31.1.91

Newton, N. (2014). 'An Aspect of Functional Harmony in Schoenberg's Early Post-Tonal Music.' *Music Analysis* 33, 1–31. https://doi.org/10.1111/musa.12018

Odria, C. (2017). 'Seeking a New Path: Pasacalle Activists Practicing Culture in Villa El Salvador, Perú.' *Ethnomusicology* 61, 1–30. https://doi.org/10.5406/ethnomusicology.61.1.0001

O'Leary, J. (2014). 'Oklahoma!, "Lousy Publicity," and the Politics of Formal Integration in the American Musical Theater.' *The Journal of Musicology* 31, 139. https://doi.org/10.1525/jm.2014.31.1.139

Park, J. (2017). 'In Pursuit of Cyclic Form: Eisler's Palmström, Op. 5.' *Journal of Musicological Research* 36, 143–169. https://doi.org/10.1080/01411896.2017.1305847

Paynter, T. (2015). 'Form and Process in Morton Feldman's Spring of Chosroes.' *Music Analysis* 34, 47–90. https://doi.org/10.1111/musa.12040

Pedneault-Deslauriers, J. (2015). 'Webern's Angels.' *The Journal of Musicology* 32, 78. https://doi.org/10.1525/jm.2015.32.1.78

Placanica, F. (2018). 'The Unsung One: The Performer's Voice in Twentieth-century Musical Monodrama.' *Journal of Musicological Research* 37, 119–140. https://doi.org/10.1080/01411896.2017.1381522

Poole, A. (2015). 'Determining Playback Speeds of Early Ethnographic Wax Cylinder Recordings: Technical and

Methodological Issues.' *Ethnomusicology Forum* 24, 73–101. https://doi.org/10.1080/17411912.2015.1010557

Purciello, M.A. (2018). 'Early Modern Comedy and the Politics of Religion: Reconsidering Comedy in Rospigliosi's Il Sant'Alessio.' *Journal of the Royal Musical Association* 143, 51–87. https://doi.org/10.1080/02690403.2018.1434341

Quintero, M. B. (2018). 'Exchange, materiality and aesthetics in Colombian champeta.' *Ethnomusicology Forum* 27, 3–24. https://doi.org/10.1080/17411912.2018.1454842

Reece, F. (2018). 'Composing Authority in Six Forged "Haydn" Sonatas.' *The Journal of Musicology* 35, 104. https://doi.org/10.1525/jm.2018.35.1.104

Rees, H. (2016). 'Environmental Crisis, Culture Loss, and a New Musical Aesthetic: China's "Original Ecology Folksongs" In Theory and Practice.' *Ethnomusicology* 60, 53–88. https://doi.org/10.5406/ethnomusicology.60.1.0053

Riley, M. (2015). 'The Sonata Principle Reformulated for Haydn Post-1770 and a Typology of his Recapitulatory Strategies.'

Journal of the Royal Musical Association 140, 1–39. https://doi.org/10.1080/02690403.2015.1008862

Rodger, G. M. (2015). 'When Singing Was Acting: Song and Character in Variety Theater.' *The Musical Quarterly* 98, 57–80. https://doi.org/10.1093/musqtl/gdv006

Rodgers, S. (2017). 'Song and the Music of Poetry.' *Music Analysis* 36, 315–349. https://doi.org/10.1111/musa.12091

Rodgers, S. (2014). 'Sentences with Words: Text and Theme-Type in Die schöne Müllerin.' *Music Theory Spectrum* 36, 58–85. https://doi.org/10.1093/mts/mtu007

Rodríguez, E. M. (2015). 'Transatlantic Networks in the Correspondence of Two Exiled Spanish Musicians, Julián Bautista and Adolfo Salazar.' *Journal of the Royal Musical Association* 140, 93–119. https://doi.org/10.1080/02690403.2015.1008864

Roig-Francolí, M.A. (2018). 'From Renaissance to Baroque: Tonal Structures in Tomás Luis de Victoria's Masses.' *Music Theory Spectrum* 40, 27–51. https://doi.org/10.1093/mts/mty001

438 of 460

Rone, V. (2018). 'Vatican II, Maurice Duruflé, and the Harmony of Resignation in Notre Père: Part Two.' *Journal of Musicological Research* 37, 166–180. https://doi.org/10.1080/01411896.2018.1451713

Ronyak, J. (2017). 'Meeting Barthes at Fischer-Dieskau's Mill.' *The Journal of Musicology* 34, 32. https://doi.org/10.1525/jm.2017.34.01.32

Rumph, S. (2017). 'Fauré's "Parnassian" Song Cycle: Poème d'un jour, Op. 21.' *The Musical Quarterly* 99, 89–133. https://doi.org/10.1093/musqtl/gdw007

Sala, L. L. (2014). 'Propaganda, Negotiations, and Antisemitism at the Teatro La Fenice, 1937–43: Proscription Lists and Other Unpublished Documents.' *Journal of Musicological Research* 33, 271–314. https://doi.org/10.1080/01411896.2014.954510

Politz, S. (2018). '"People of Allada, This Is Our Return": Indexicality, Multiple Temporalities, and Resonance in the Music of the Gangbé Brass Band of Benin.' *Ethnomusicology*

62, 28–57. https://doi.org/10.5406/ethnomusicology.62.1.0028

Schmalfeldt, J. (2017). 'Beethoven's "Violation": His Cadenza for the First Movement of Mozart's Concerto in D Minor, K. 466.' *Music Theory Spectrum* 39, 1–17. https://doi.org/10.1093/mts/mtx001

Schmelz, P. J. (2015). '"Shostakovich" Fights the Cold War: Reflections from Great to Small.' *Journal of Musicological Research* 34, 91–140. https://doi.org/10.1080/01411896.2015.1020250

Schmidt-Beste, T. (2014). 'Preventive and Cautionary Dynamics in the Symphonies of Mendelssohn and his Time.' *The Journal of Musicology* 31, 43. https://doi.org/10.1525/jm.2014.31.1.43

Severn, J. R. (2015). 'Beyond Falstaff in "Die lustigen Weiber von Windsor": Otto Nicolai's Revolutionary "Wives".' *Music and Letters* 96, 28–54. https://doi.org/10.1093/ml/gcv003

Sheppard, W. A. (2015). 'Puccini and the Music Boxes.' *Journal of the Royal Musical Association* 140, 41–92. https://doi.org/10.1080/02690403.2015.1008863

Simms, B. R. (2017). 'Paul Kammerer and Alban Berg: Rivals in Song.' *Music and Letters* 98, 104–127. https://doi.org/10.1093/ml/gcx002

Slim, H. C. (2018). 'Stravinsky in Vancouver, 1917–2017: Concerts, Premieres, Collections.' *The Musical Quarterly* 100, 269–296. https://doi.org/10.1093/musqtl/gdy003

Sloan, N. (2019). 'Stormy Relations: The Cotton Club, Broadway Spirituals, and Harlem Encounters in the Music of Harold Arlen and Ted Koehler.' *The Musical Quarterly* 101, 120–156. https://doi.org/10.1093/musqtl/gdy017

Slobin, M. (2016). 'Improvising a Musical Metropolis: Detroit in the 1940s–1960s.' *Ethnomusicology* 60, 1–21. https://doi.org/10.5406/ethnomusicology.60.1.0001

Smith, P. H. (2015). 'Hausmusik for Cognoscenti: Some Formal Characteristics of Schumann's Late-Period Character Pieces for

Instrumental Ensemble.' *Music Theory Spectrum* 37, 51–70. https://doi.org/10.1093/mts/mtv001

Solvik, M., & Hefling, S. E. (2014). 'Natalie Bauer-Lechner on Mahler and Women: A Newly Discovered Document.' *The Musical Quarterly* 97, 12–65. https://doi.org/10.1093/musqtl/gdu002

Sorbo, L. (2015). 'Alfredo Casella's War Pages.' *Journal of Musicological Research* 34, 51–70. https://doi.org/10.1080/01411896.2015.987632

Steege, B. (2017). 'Antipsychologism in Interwar Musical Thought: Two Ways of Hearing Debussy.' *Music and Letters* 98, 74–103. https://doi.org/10.1093/ml/gcx011

Stoecker, P. (2014). 'Aligned Cycles in Thomas Adès's Piano Quintet.' *Music Analysis* 33, 32–64. https://doi.org/10.1111/musa.12019

Stoessel, J. (2014). 'Arms, A Saint and Inperial sedendo fra più stelle.' *The Journal of Musicology* 31, 1. https://doi.org/10.1525/jm.2014.31.1.1

Stoia, N. (2015). 'Triple Counterpoint and Six-Four Chords in Bach's Sinfonia in F Minor.' *Music Analysis* 34, 305–334. https://doi.org/10.1111/musa.12041

Straus, J. N. (2014). 'Harmony and Voice Leading in the Music of Stravinsky.' *Music Theory Spectrum* 36, 1–33. https://doi.org/10.1093/mts/mtu008

Strunk, S. (2016). 'Tonal and Transformational Approaches to Chick Corea's Compositions of the 1960s.' *Music Theory Spectrum* 38, 16–36. https://doi.org/10.1093/mts/mtv037

Strykowski, D. R. (2016). 'The Diegetic Music of Berg's Lulu: When Opera and Serialism Collide.' *Journal of Musicological Research* 35, 1–22. https://doi.org/10.1080/01411896.2015.1082066

Summers, T. (2015). '"Sparks of Meaning": Comics, Music and Alan Moore.' *Journal of the Royal Musical Association* 140, 121–162. https://doi.org/10.1080/02690403.2015.1008865

Tarsi, B. (2017). 'At the Intersection of Music Theory and Ideology: A. Z. Idelsohn and the Ashkenazi Prayer Mode Magen Avot.'

Journal of Musicological Research 36, 208–233. https://doi.org/10.1080/01411896.2017.1340033

Taylor, B. (2016). 'Mutual Deformity: Ignaz Moscheles's Seventh and William Sterndale Bennett's Fourth Piano Concertos.' *Music Analysis* 35, 75–109. https://doi.org/10.1111/musa.12062

Taylor, B. (2014). 'Schubert and the Construction of Memory: The String Quartet in A minor, D.804 ('Rosamunde').' *Journal of the Royal Musical Association* 139, 41–88. https://doi.org/10.1080/02690403.2014.886414

Tenzer, M. (2015). 'Meditations on Objective Aesthetics in World Music.' *Ethnomusicology* 59, 1–30. https://doi.org/10.5406/ethnomusicology.59.1.0001

Terrigno, L. (2017). 'Tonal Problems as Agents of Narrative in Brahms's Unbewegte laue Luft, Op. 57 No. 8.' *Music Analysis* 36, 350–371. https://doi.org/10.1111/musa.12100

Treitler, L. (2017). 'Max Raabe in Israel: Living Memorializing.' *The Musical Quarterly* 99, 81–88. https://doi.org/10.1093/musqtl/gdw012

Trippett, D. (2018). 'Music and the Transhuman Ear: Ultrasonics, Material Bodies, and the Limits of Sensation.' *The Musical Quarterly* 100, 199–261. https://doi.org/10.1093/musqtl/gdy001

Tsai, T.-H. (2018). 'Music as Queering in E. M. Forster's "Goldsworthy Lowes Dickinson."' *Music and Letters* 99, 1–15. https://doi.org/10.1093/ml/gcx093

Utz, C. (2017). 'Time-Space Experience in Works for Solo Cello by Lachenmann, Xenakis and Ferneyhough: a Performance-Sensitive Approach to Morphosyntactic Musical Analysis.' *Music Analysis* 36, 216–256. https://doi.org/10.1111/musa.12076

Vande Moortele, S. (2017). 'Murder, Trauma, and the Half-Diminished Seventh Chord in Schoenberg's "Song of the Wood Dove."' *Music Theory Spectrum* 39, 66–82. https://doi.org/10.1093/mts/mtx005

Venn, E. (2014). 'Thomas Adès's "Freaky, Funky Rave."' *Music Analysis* 33, 65–98. https://doi.org/10.1111/musa.12020

Walden, J. S. (2014). 'The "Yidishe Paganini": Sholem Aleichem's Stempenyu, the Music of Yiddish Theatre and the Character of the Shtetl Fiddler.' *Journal of the Royal Musical Association* 139, 89–136. https://doi.org/10.1080/02690403.2014.886428

Waltham-Smith, N. (2018). 'Form and Repetition: Deleuze, Guillaume and Sonata Theory.' *Music Analysis* 37, 150–183. https://doi.org/10.1111/musa.12098

Waltham-Smith, N. (2017). 'The Time It Takes to Listen.' *Music Theory Spectrum* 39, 18–35. https://doi.org/10.1093/mts/mtx006

Waters, K. (2016). 'Chick Corea and Postbop Harmony.' *Music Theory Spectrum* 38, 37–57. https://doi.org/10.1093/mts/mtv024

Weaver, A. H. (2017). 'Memories Spoken and Unspoken: Hearing the Narrative Voice in Dichterliebe.' *Journal of the Royal Musical Association* 142, 31–67. https://doi.org/10.1080/02690403.2017.1286123

Weaver, A. M. (2016). 'The Soprano and the Nightingale: Aleksandr Alyabyev's "Solovey."' *Journal of Musicological Research* 35, 23–44. https://doi.org/10.1080/01411896.2016.1116349

Webster-Kogen, I. (2014). 'Song Style as Strategy: Nationalism, Cosmopolitanism and Citizenship in The Idan Raichel Project's Ethiopian-influenced Songs.' *Ethnomusicology Forum* 23, 27–48. https://doi.org/10.1080/17411912.2013.879034

Weitz, S. B. (2018). 'Monochromatic and Polychromatic Performance: Improvisatory Alteration in Early Nineteenth-Century French Pianism.' *The Musical Quarterly* 101, 76–109. https://doi.org/10.1093/musqtl/gdy011

Wiens, K. (2014). 'A Serbian Jewish Perspective on Media and Musical Transmission.' *Ethnomusicology Forum* 23, 49–66. https://doi.org/10.1080/17411912.2013.875733

Witulski, C. (2018). 'Contentious Spectacle: Negotiated Authenticity within Morocco's Gnawa Ritual.' *Ethnomusicology* 62, 58–82. https://doi.org/10.5406/ethnomusicology.62.1.0058

Woodfield, I. (2018). '"Che soave zeffiretto" and the Structure of Act 3 of Le nozze di Figaro.' *Journal of the Royal Musical Association* 143, 89–136. https://doi.org/10.1080/02690403.2018.1434345

Ying-fen, W. (2016). 'The Transborder Dissemination of Nanguan in the Hokkien Quadrangle before and after 1945.' *Ethnomusicology Forum* 25, 58–85. https://doi.org/10.1080/17411912.2016.1151816

Zanovello, G. (2016). '"You Will Take This Sacred Book": The Musical Strambotto as a Learned Gift.' *Journal of the Royal Musical Association* 141, 1–26. https://doi.org/10.1080/02690403.2016.1151230

Appendix 1: SCImago Journal Rankings (Top 100 Journals)

RANK	Title	SJR	H Index	Total Refs.	Total Cites (3years)	Country
1	IEEE Signal Processing Magazine	1,364	155	3093	1969	U.S
2	Journal of Research in Music Education	987	37	1202	97	U.S
3	Psychology of Music	980	51	4930	380	UK
4	Empirical Studies of the Arts	814	12	840	45	U.S
5	Musicae Scientiae	764	31	2261	129	U.S.
6	Music Perception	684	54	2240	162	U.S.
7	Journal of Aesthetics and Art Criticism	577	27	1144	76	U.S.
8	Research Studies in Music Education	541	22	1269	38	UK
9	Journal of Music Teacher Education	526	15	835	53	U.S.
10	Music Education Research	507	16	2000	92	UK
11	British Journal of Music Education	499	19	890	52	UK
12	Journal of Music Therapy	474	43	746	119	U.S.
13	Journal of African Cultural Studies	455	10	1585	85	UK
14	International Journal of Music Education	452	23	2111	80	U.S.
15	Notes	417	8	410	14	U.S.
16	Music Therapy Perspectives	382	7	1179	38	U.S.
17	Acta Acustica united with Acustica	379	49	3520	399	Germany
18	Journal of New Music Research	368	23	1469	90	UK
19	International Journal of Community Music	362	7	569	40	UK
20	AES: Journal of the Audio Engineering Society	313	55	2126	309	U.S.

21	Popular Music and Society	304	21	1839	63	UK
22	Journal of Mathematics and Music	261	11	183	24	U.S.
23	Music Reference Services Quarterly	260	7	186	13	U.S.
24	Revista Electronica Complutense de Investigacion Musical	252	4	432	11	Spain
25	Bulletin of the Council for Research in Music Education	238	14	802	22	U.S.
26	Computer Music Journal	236	36	472	59	U.S.
27	Organised Sound	226	25	1147	43	UK
28	Music Theory Spectrum	225	19	806	21	U.S.
29	Ethnomusicology	194	17	849	31	U.S.
30	Leonardo	192	20	1522	54	U.S.
31	Music Scholarship	187	4	1120	61	Russia
32	International Review of the Aesthetics and Sociology of Music	179	3	438	2	Croatia
33	Journal of Music Theory	173	12	413	9	U.S.
34	American Music	172	9	1316	12	U.S.
35	Journal of Musicology	162	14	1223	12	U.S.
36	Ethnomusicology Forum	160	4	776	15	UK
37	Journal of the American Musicological Society	158	22	1359	15	U.S.
38	Traditiones	157	6	87	20	Slovenia
39	Twentieth-Century Music	151	12	875	14	UK
40	Folk Music Journal	148	5	575	2	UK
41	Popular Music	147	30	1091	17	UK

42	Journal of Folklore Research	143	13	815	17	U.S.
43	Early Music	142	14	1663	19	UK
44	American Imago	139	11	562	5	U.S.
45	Journal of the Royal Musical Association	139	14	415	7	UK
46	Modernism/ Modernity	138	19	1612	19	U.S.
47	Musical Quarterly	135	15	494	12	U.S.
48	Journal of Musicological Research	131	11	1018	13	UK
49	Journal of the Society for American Music	131	10	682	14	UK
50	Yearbook for Traditional Music	131	9	0	11	U.S.
51	Leonardo Music Journal	126	10	276	8	U.S.
52	Muziki	124	5	453	5	U.S.
53	Contemporary Music Review	123	15	1112	16	UK
54	Fontes Artis Musicae	123	4	286	5	Switz-er-land
55	Nineteenth century Music	123	15	435	4	U.S.
56	Music Analysis	121	16	538	6	UK
57	Opus	119	1	900	2	Brazil
58	I Tatti Studies	116	5	1060	3	Italy
59	Journal of Music, Technology and Education	116	9	239	18	UK
60	Journal of Popular Music Studies	116	10	0	18	Den-mark
61	Music and Letters	116	13	487	6	UK
62	Musicology Australia	113	7	218	2	UK
63	Oido Pensante	112	1	376	1	Ar-gen-tina
64	Anuario Musical	111	2	861	5	Spain
65	Musica Hodie	111	3	666	4	Brazil

66	Music Theory Online	111	2	1590	5	U.S.
67	Opera Quarterly	111	9	531	4	U.S.
68	Resonancias	111	1	329	3	Chile
69	Muzikoloski Zbornik	110	4	542	2	Slo-venia
70	Arti Musices	106	1	271	3	Croa-tia
71	Co-herencia	106	3	1097	4	Co-lom-bia
72	Eighteenth-Century Music	106	9	306	3	UK
73	Journal of the American Musical Instrument Society	105	2	0	2	U.S.
74	Malaysian Music Journal	105	1	0	5	Malay-sia
75	Tempo	105	7	298	8	UK
76	MusikTheorie	104	3	617	0	Ger-many
77	Acta Ethnographica Hungarica	103	5	567	7	Hun-gary
78	Jazz Research Journal	103	0	0	0	UK
79	Journal of Band Research	103	6	220	0	U.S.
80	Black Music Research Journal	102	7	0	1	U.S.
81	Musik in Bayern	102	1	0	0	Ger-many
82	Perspectives of New Music	102	2	0	4	U.S.
83	Acta Musicologica	101	6	388	5	Switz-er-land
84	Art Inquiry	101	1	284	0	Po-land
85	Bach	101	2	292	2	U.S.
86	Cuadernos de Musica, Artes Visuales y Artes Escenicas	101	2	1353	4	Co-lom-bia
87	Early Music History	101	11	419	0	UK

88	Jazz Perspectives	101	3	278	2	UK
89	Journal of Science and Technology of the Arts	101	1	138	1	Portu-gal
90	Journal of the Musical Arts in Africa	101	4	323	0	UK
91	Journal of World Popular Music	101	2	517	5	UK
92	Musicologica Brunensia	101	0	327	0	Czech Re-public
93	Musikforschung	101	3	302	0	Ger-many
94	Opera	101	1	0	0	UK
95	PAJ - Journal of Performance and Art	101	9	97	3	U.S.
96	Perfect Beat	101	8	181	1	UK
97	Plainsong and Medieval Music	101	8	141	2	UK
98	Popular Entertainment Studies	101	2	321	4	Aus-tralia
99	Revista de Musicologia	101	0	739	0	Spain
100	Tijdschrift van de Koninklijke Vereniging voor Nederlandse Muz	101	4	209	1	Neth-er-lands

Appendix 2: Journal Statements of Purpose

28 Music Theory Spectrum

https://societymusictheory.org/music-theory-spectrum

A leading journal in the field and an official publication of the Society for Music Theory, Music Theory Spectrum features articles on a wide range of topics in music theory and analysis, including aesthetics, critical theory and hermeneutics, history of theory, post-tonal theory, linear analysis, rhythm, music cognition, and the analysis of popular musics. The journal welcomes interdisciplinary articles revealing intersections with topics in other fields such as ethnomusicology, mathematics, musicology, philosophy, psychology, and performance.

29 Ethnomusicology

https://www.ethnomusicology.org/page/Pub_Journal

Ethnomusicology is an international, peer-reviewed journal published three times a year by the University of Illinois Press on behalf of the Society for Ethnomusicology. It features scholarly articles representing theoretical perspectives and research in ethnomusicology and related fields, as well as book, sound recording, film, video, and multimedia reviews.

35 Journal of Musicology

https://jm.ucpress.edu/

The *Journal of Musicology* (*JM*) is a refereed, international quarterly journal devoted to exemplary scholarship across the spectrum of music studies. As a leading disciplinary forum founded in 1981 by Marian Green, *JM* reflects the breadth of musicology today and helps to shape the future trajectory of cross-disciplinary humanistic inquiry. Through its long-standing tradition of publishing the most exciting work by younger scholars alongside the contributions of senior scholars, *JM* continues to expand the perimeters of musicology in a spirit of openness, diversity, and academic excellence.

36 Ethnomusicology Forum

https://www.tandfonline.com/action/journalInformation?show=aimsScope&journalCode=remf20

Ethnomusicology Forum, formerly known as the British Journal of Ethnomusicology, is the academic, refereed journal of the British Forum for Ethnomusicology. The journal seeks to provide a dynamic forum for the presentation of new thinking in the field of ethnomusicology, defined broadly as the study of "people making

455 of 460

music", and encompasses the study of all music, including Western art music and popular music.

Articles often emphasize first-hand, sustained engagement with people as music makers, taking the form of ethnographic writing following one or more periods of fieldwork. Typically, ethnographies aim for a broad assessment of the processes and contexts through and within which music is imagined, discussed and made. Ethnography may be synthesised with a variety of analytical, historical and other methodologies, often entering into dialogue with other disciplinary areas such as music psychology, music education, historical musicology, performance studies, critical theory, dance, folklore and linguistics. The field is therefore characterised by its breadth in theory and method, its interdisciplinary nature and its global perspective.

45 Journal of the Royal Musical Association

https://www.tandfonline.com/action/aboutThisJournal?
show=aimsScope&journalCode=rrma20

The **Journal of the Royal Musical Association** was established in 1986 (replacing the Association's Proceedings) and is now one of the major international refereed journals in its field. Its editorial policy is to publish outstanding articles in fields ranging from

historical and critical musicology to theory and analysis, ethnomusicology, and popular music studies. The journal works to disseminate knowledge across the discipline and communicate specialist perspectives to a broad readership, while maintaining the highest scholarly standards.

Research articles published in the journal have undergone rigorous double-blind peer-review after initial screening by the Editor for quality and fit with the journal's objectives.

47 Musical Quarterly

https://academic.oup.com/mq/pages/About

The Musical Quarterly, founded in 1915 by Oscar Sonneck, has long been cited as the premier scholarly musical journal in the United States. Over the years it has published the writings of many important composers and musicologists, including Aaron Copland, Arnold Schoenberg, Marc Blitzstein, Henry Cowell, and Camille Saint-Saens. The journal focuses on the merging areas in scholarship where much of the challenging new work in the study of music is being produced.

Regular sections include 'American Musics', 'Music and Culture', 'The Twentieth century and Beyond', and an 'Institutions, Industries,

Technologies' section which examines music and the ways it is created and consumed. In addition, a fifth section entitled 'Primary Sources' features discussions on issues of biography, texts, and manuscripts; reflections on leading figures; personal statements by noted performers and composers; and essays on performances and recordings. Along with discussions of important new books, *MQ* also publishes review essays on a wide variety of significant new music performances and recordings.

48 Journal of Musicological Research

https://www.tandfonline.com/action/journalInformation?show=aimsScope&journalCode=gmur20

The **Journal of Musicological Research** publishes original articles on all aspects of the discipline of music: historical musicology, style and repertory studies, music theory, ethnomusicology, music education, organology, and interdisciplinary studies.

Because contemporary music scholarship addresses critical and analytical issues from a multiplicity of viewpoints, the **Journal of Musicological Research** seeks to present studies from all perspectives, using the full spectrum of methodologies. This variety makes the Journal a place where scholarly approaches can coexist,

in all their harmony and occasional discord, and one that is not allied with any particular school or viewpoint.

56 Music Analysis

https://onlinelibrary.wiley.com/page/journal/14682249/homepage/ productinformation.html

Overview

Music Analysis is the international forum for the presentation of new writing focused on musical works and repertoires. Through articles of this kind and through its lively Critical Forum, it also aims to take forward debates concerning the relationship of technical commentary on music with music theory, critical theory, music history and the cognitive sciences.

Music Analysis is eclectic in its coverage of music from medieval to post-modern times, and has regular articles on non-western music. Its lively tone and focus on specific works makes it of interest to the general reader as well as the specialist.

Aims and Scope

Founded in 1982, *Music Analysis* publishes major orientation articles by respected scholars such as Kofi Agawu, Craig Ayrey, Richard Cohn, Nicholas Cook, Hermann Danuser and Marianne Kielian-

Gilbert.	The journal has also featured translations of important articles by Adorno, Molino, Ratz, Ruwet and Schenker.

Appendix 3: Novice Information

| Country | Institution | Secondary | | Tertiary | | |
| | | | | Bachelor's | | Master's |
		#	Year	#	Year	#
China	Beijing World Youth Academy	4	3			
Australia	University of Melbourne			2	2	
UK	Goldsmiths, University of London			2	1	
	Guildhall School			2	1	1
	Royal Northern College of			2	1	
USA	Baylor University					1