UNIVERSITY OF CALIFORNIA,
IRVINE

Conversational Composition
Techniques of composition derived from and driven by conversation

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Music
Integrated Composition, Improvisation, and Technology

by

Tomoko Ozawa

Dissertation Committee:
Professor Chris Dobrian, Chair
UCI Chancellor's Professor Kei Akagi
Professor John Crawford
Professor Michael Dessen
Professor Kojiro Umezaki

2021

DEDICATION

I dedicate this dissertation to my parents who have continuously supported myself and my work from afar with complete faith and trust in everything I do. I would like to thank the support and encouragement of Ruth McKay and Don Campbell without whom I would not have been able pursue my PhD degree.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to express the sincerest appreciation and respect to my committee chair Chris Dobrian who has guided me in every step towards completing this dissertation with exhaustive patience, compassion and an open mind. I would like to thank my committee members, UCI Chancellor's Professor Kei Akagi, Professor John Crawford, Professor Michael Dessen and Professor Kojiro Umezaki, for their wisdom, insights, tremendous support and encouragement. I would like to thank Professor Mari Kimura for introducing me to the motion sensor MUGIC™ and granting me permission to use it in my piece, as well as coaching me in computer music and writing for strings.

In addition, I would like to thank and acknowledge the immense support from my musicians: Bella Pepke, Blake Harrison-Lane and JoVia Armstrong whom I am honored to call my friends. Thank you for teaching me and sharing your gifts with me. Special thanks to Teerath Majumder for filming and streaming the concert, Bruce Warner for his generous support during my xMPL residency, and Diana Herrera for the lighting works.

Lastly, I would like to thank the Claire Trevor School of the Arts for funding my studies and giving me permission to use the Disklavier and the recording studios for this

research. This dissertation would not have been possible without all of your support. I am deeply grateful for such an invaluable learning experience I had at UC Irvine.

I would also like to express my gratitude to my former teachers and mentors who taught me the joy of creating music in jazz, free jazz, classical, contemporary classical, computer music, etc. All of that has inspired me to be who I am as a musician today.

VITA

Tomoko Ozawa

2007	B.A. in Liberal Arts, International Christian University, Tokyo
2007-10	Software Technical Support Engineer, IBM Japan, Tokyo
2011-14	Berklee College of Music, Boston
2014	Diploma with focus in jazz composition, Berklee College of Music
2014-2017	Piano Accompanist, Boston Conservatory and Boston Ballet School, Boston
2015-17	Longy School of Music of Bard College, Cambridge
2017	M.M. in Piano Performance in Modern American Music, Longy School of Music
2018-2021	Teaching Assistant, University of California, Irvine
2021	Ph.D. in Music, University of California, Irvine

ABSTRACT

Conversational Composition

Techniques of composition derived from and driven by conversation

by

Tomoko Ozawa

Doctor of Philosophy in Music

University of California, Irvine, 2021

Professor Chris Dobrian, Chair

The metaphor of conversation has often been used with regard to interaction between instruments or sections of instruments in an ensemble. Interaction is a common practice in jazz and other improvised music, while in Western classical music, interactivity is considered implicit in the score. However, the conversational aspect of the music is often left as merely a metaphor, without that comparison being fully explored. In this dissertation, compositional techniques derived from theoretical models of conversation—such as turn-taking, stages of engagement and disengagement, and types of interruption during conversation—are demonstrated through my recent compositions for an ensemble and for solo piano. The intent is not only to explore alternative ways of composition, but also to offer a composing-performance practice that invites performers to be more actively involved in the process of creating music as a collaborative work,

while allowing them to fully express their individualities and communicate with each other through music. Hence, the performance itself becomes the manifestation of conversation.

This dissertation consists of original compositions along with a written thesis. The original compositions include a series of ensemble pieces, *Group Talk No. 1-3* for a quartet performed by Blake Harrison-Lane on violin, Bella Pepke on cello, JoVia Armstrong on percussion and myself on piano, as well as solo works for Disklavier, *Phasing* and *Streams of Talk* and a piece for piano and motion sensor MUGIC™, *Cricket Wind*. The written thesis includes four main chapters. Chapter 1 reviews theories of conversation such as turn-taking models and different phases and stages of conversation. Chapter 2 surveys other composers' compositional and improvisational approaches, examining scores and concepts that are relevant to topics of conversation. Chapter 3 demonstrates how I translated conversational models into music, using examples from my own musical works. Chapter 4 discusses the rehearsal process and the concert performance of these works, to examine what worked or did not work as expected in terms of the initial purpose of this research, and some lessons learned from this experience.

Introduction

The primary purpose of this dissertation is to seek an answer to my ongoing questions: How can I best compose music which represents a conversation, and how can I compose in a way that encourages performers to interact? Interactions in music have often been discussed using the analogy of conversation. However, such an analogy of conversation is merely metaphoric, which only suggests any kind of interaction between instruments, without referring to specific theories of conversation. It is also common for such conversational behavior to be realized only after the performance. My approach, on the other hand, is to investigate how the practice of conversation can be used as a compositional device, and this dissertation therefore surveys my own musical works which implement conversational models in composition and improvisation.

When jazz musicians perform, a soloist will create a melodic or rhythmic motif, then the band may pick it up. The members of the rhythm section might respond in several different ways; they can simply mimic the idea somewhere in the bass line or in the rhythm of the drums, or inject something in between statements of the motif, as if the rhythm section is responding to the soloist. If the soloist changes the time feel, the band may adjust the time feel to match or they may choose to be counteractive by remaining in the previous time feel. These common behaviors in jazz are often discussed using an analogy of conversation. That analogy is theorized in depth by Ingrid Monson in her book *Saying Something: Jazz Improvisation and Interaction*, based on interviews she conducted with masters of jazz. As one of Monson's subjects describes it, jazz-based improvisational music making is “like a conversation and one guy will … create a melodic

motif or a rhythmic motif and the band picks it up. It's like saying that you all are talking about the same thing," (Monson 1996, 31). Similar ideas are also discussed in another prominent ethnographic study of jazz by Paul Berliner, who uses the "metaphor of conversation to describe aspects of the improvisational process," and describes "jazz as a musical language, improvisation as musical conversation, and good improvisation as talking or 'saying something'" (Berliner 1994, 74). Preeminent American musicologist, Charles Seeger, was also aware of the relationship between conversation and music, and considered music and language as subsets of aural communication, writing that "if improvisation is like conversation (subset of music and language), then sociable, face-to-face communication (subset of communication) may be the larger category at stake" (Quoted in Monson, 76).

Beyond being an interesting metaphor for formal structures in jazz, conversation carries the idea of performers having freedom and agency to express themselves as individuals while still working collectively. It also takes on particularly powerful meanings when understood through the lens of African American history and culture. The tradition of jazz has been an ever-evolving form of artistic expression since its inception, constantly challenging musical "rules" and norms, breaking and merging genres and joining people of different colors and cultures.

The thesis of this set of my recent compositions is to take the conversational aspect of improvisation literally as a model or technique for composition. Based on theorized conversational models, I have composed a suite for an ensemble, and three solo pieces in which the instrumental soloist engages in a dialogic relationship with a computer program. In the ensemble works, the conversational models are utilized to shape and advance the music for musicians who are trained in different musical practices, specifically those trained primarily in

jazz and those trained in Western classical music. In this regard, my plan—to develop new skills and capacities within performers by increasing the sense of agency—extends beyond acquisition of new skills on their instruments. The intent is to grant the agency for self-expression, encouraging performers to engage in the collective work, and to question and develop ideas as a group. The objective is to introduce them to new musical and cultural value systems, new sensibilities and attitudes about creating music, and new ways of conceiving their own role as a musician. In my solo works, I am using interactive programming to create imagined conversations between myself and other persons or beings in nature, depicting different phases of conversation.

In musical contexts, communication among performers or between performers and conductor is performed through a combination of sounds and gestures, which are used to advance the music or create some sort of change. Those sounds and gestures used for advancing or altering the music are called cues. Performers will look out for the cues to determine whether or not to take certain actions by vigorously listening to or looking at each other or at the conductor. While various types of cues, both gestural[1] and sonic, have been developed and employed to change or advance the music, most commonly in jazz and other kinds of improvised music, in this research I will frame the focus around sonic cues instead of gestural cues. Sonic cues are sound characteristics which indicate performers to take certain actions when established during the performance. Hence, they can be considered as a codified language which can symbolize a question or statement, and responses to be exchanged between the performers. One primary reason why I choose sonic cues over gestural cues is that this research models aural

[1] Conducting is the most common use of gestural cues. Butch Morris has developed a comprehensive vocabulary of conducting gestures which he employs in an improvising orchestra (Morris et al. 2017).

conversation, in which the interaction occurs directly between the individuals in the group without having a mediator or a director in the way. Another reason is that sonic cues can be implemented directly in the music, and are used in such a way that the performers have the agency to take initiative. This idea stems from the idea of "conversational" improvisational process in jazz. In this light, compositions based on conversational models using sonic cues will allow the performers to initiate cues more organically during the course of performance without having to take their hands from their instruments, solely dependent on aural communication.

Therefore, the primary objective of this research has been to develop compositional techniques derived from the practice of conversation, which encourages performers to interact. Chapter 1 will discuss theories of conversation such as turn-taking models and different phases and stages of conversation. Chapter 2 will survey other composers' compositional and improvisational approaches, examining scores and concepts that are relevant to topics of conversation. In chapter 3, I will explain how I translated conversational models into music, using examples from my own musical works. Chapter 4 will discuss the rehearsal process and the concert performance of these works, to examine what worked or did not work as expected in terms of the initial purpose of this research, and some lessons learned from this experience.

Chapter 1: Conversation Analysis

Overview of conversation analysis

Conversation Analysis (CA) is a field of study that examines the norms, practices and competencies constructing the organization of social interaction. Forms of spoken interaction addressed in CA include everyday conversations between friends and acquaintances, extending to interactions in medical, educational, mass media and socio-legal contexts, interactions such as lecturing or speech-making, and technologically complex interactions such as web-based multiparty communication. Regardless of the interaction being studied, CA starts from the perspective that the details of conduct in interaction are highly organized and orderly. In addition, it stands on the assumption that the specificities of meaning and understanding in interaction would only function with this orderliness, as Drew and Heritage state in the introduction of their comprehensive collection of works, *Conversation Analysis* (2006).

As Jack Sidnell explains, CA began to develop as a field in the 1960s through the collaboration of Harvey Sacks, Emanuel Schegloff, and Gail Jefferson, and it was particularly indebted to the sociologists Erving Goffman and Harold Garfinkel (Sidnell 2011, 6). In particular, Sidnell points to an early contribution by Goffman who, in contrast to most prevailing researchers of the time, "was very clear on the point that interaction had properties specific to it and had to be studied on its own terms" (Sidnell, 7). In one of his early publications *Alienation from Interaction* (1957), Goffman mentioned that “[w]e must also see that a conversation has a life of its own and makes demands on its behalf. It is a little social system with its own boundary-maintaining tendencies” (Goffman 1957, 47). In the article *Interaction Order* (1983),

he suggested that the organization of human interaction, in his words "the interaction order" (Goffman 1983), established its own social institution, and that it is itself a moral ordering involving a complex web of standards, expectations, rules and proscriptions. He considered face-to-face interaction as its own institution and, at the same time, the fundamental element for social organizations (Goffman 1983, 7).

In the late 1950's and early 1960's, Harold Garfinkel was also developing a critique of mainstream sociological thinking that was to develop into ethnomethodology. For instance, Garfinkel challenged the conventional wisdom promoted by Talcott Parsons with whom he studied in the social relations program at Harvard—"social order is a result of socialization and the internalization of norms"—by suggesting that, "to the extent that social life is regulated by such norms, this rests upon a foundation of practical reasoning" (Sidnell, 8). Instead of ascribing the meaning of an utterance to conventional rules or norms, Garfinkel related the action and how its decision was made to the context within which it was produced.

As Sidnell explains, Goffman and Garfinkel's work during this period represent the two contrasting intellectual streams in sociology that are often considered to have led to the emergence of CA. However, there were other researchers such as Sacks, Schegloff and Jefferson who were developing their own approaches to CA, influenced by but also distinct from the works of Goffman and Garfinkel. By the late 1960's, Sacks was already demonstrating a unique approach to CA examining police assessment of moral character (Sacks, Schegloff and Jefferson 1974, 280-293). S*equencing in conversational openings* (Schegloff 1968) by Schegloff is also considered a model of CA method and analysis, and shows only minuscule resemblance to the studies of Goffman and Garfinkel (Sidnell, 9). In addition, there were also other, less prominent

yet important contributors to the development of CA from anthropological streams in the social sciences. Evans Pritchard's Witchcraft, *Oracles and Magic among the Azande* (1937) was another influence of Sacks, in which he saw relevance to the analysis of conversation in everyday activities of the Azandes (Sidnell, 9). In the 1960's, there were also a number of researchers working to expand "cognitive revolution" the field of anthropology, accelerated by works of Noam Chomsky such as *Aspects of a Theory of Syntax* (1964) and *Syntactic Structure* (1957) in linguistics. According to Sidnell, both Sacks and Garfinkel were influenced by these anthropological approaches; Garfinkel established "ethnomethodology"[2] which was adapted from ethnoscience (Garfinkel 1974). In *Membership Categorization Devices*, Sacks shows his interests in categories likely to be influenced by the same strands of ethnomethodology (Sacks 1989). Beside Sacks and Garfinkel, there were other researchers such as Dell Hymes and John Gumperz who contributed to the development of an approach known as ethnography of speaking, which involved describing the particular ways in which people used language as part of a distinctive cultural repertoire (Gumperz and Hymes 1964).

Another figure mentioned in Sacks's lectures (Sacks 1995) and Schegloff's writing (Schegloff 1963) in his early period was Sigmund Freud who theorized psychoanalysis. As an astute observer of human behavior, Freud contributed to educating people on the importance of behaviors that seemed initially inconsequential (Sidnell, 10). In a similar way, Sacks makes a remark in his own studies of "small" phenomena as he mentions "[n]ow such a view tends to be heavily controlled by an overriding interest in what are the first instance known to be 'big issues,' and not those which are terribly mundane, occasional, local, and the like" (Sacks 1984, 24).

[2] Ethnomethodology is a methodology for studying social life originated by Garfinkel who criticized conventional social theory (Pillay 2019).

Ultimately, CA came into being as a distinct field "in dialogue with a range of perspectives within the social and human sciences" having intimate connection with anthropology from its earliest days (Sidnell, 11).

In *Conversation Analysis* (2006)—a collection of the key articles in CA since the 1960's —Paul Drew and John Heritage outline the basic background and mechanism of CA. They analyze the foundations of the fundamental dimensions of human sociality, grouping the characteristics of those foundations in four categories: 1) turn-taking and repair, 2) sequence organization, 3) turn design and action formation and 4) institutional interaction. The chart Chronology of Conversation Analysis in Figure 1.1 (Drew and Heritage 2006) shows publications in these four categories from 1970 through the early 2000s. We can see that the approach to CA has become more diverse since 1970's, when the only publications seen were topics around Turn-taking and Repair, including the most prominent work by Sacks, Schegloff and Jefferson (Sacks, Schegloff and Jefferson 1974; 1977). Between the mid 1970's and 1980's, other areas of focus within CA began to emerge such as sequence organization, turn design, and action formation. From the 1980's onward, the topics of CA continued to branch out further. In particular, three publications between 2000 and 2005 were focused on institutional interactions, studies of CA conducted in highly contextualized settings such as critical medical visits and their implications for patients' participation (Robinson 2003), and press conferences of U.S. presidents Eisenhower and Reagan (Clayman and Heritage 2002).

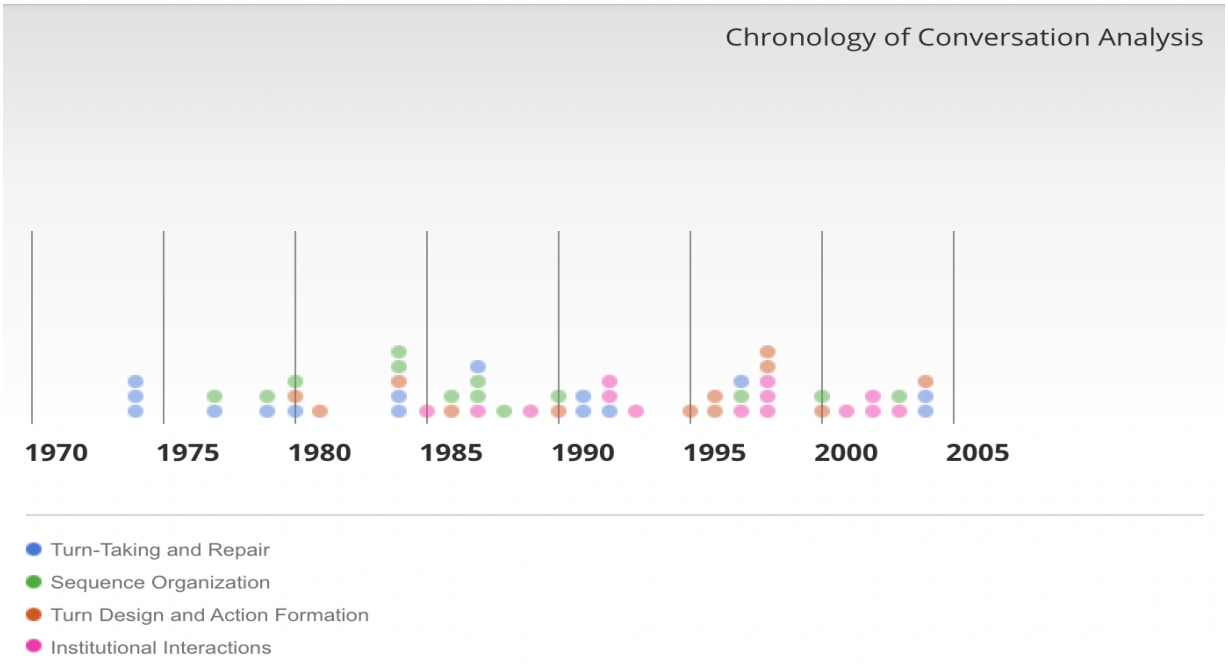


Figure 1.1 Chronology of Conversation Analysis

In the updated edition *Contemporary Studies in Conversational Analysis* (2013), Drew and Heritage have included a variety of institutional or task-focused interactions that have emerged after the publication of the first volumes, in parallel to existing domains of CA. Drew and Heritage claim that more recent research has evolved organically out of the foundational domains mentioned above, revealing both continuities and discontinuities with the classical agenda. In this way, CA research has grown exponentially, connecting with other areas in sociology, anthropology, communication, linguistics, pragmatics, and sociolinguistics (Sidnell; Stivers, 2014). Articles in this book are categorized in four domains including the three most prominent themes which stemmed from the classical tradition in CA such as social action and turn design (Stivers 2010; Pillet-Shore 2011); recipiency and responses (Holt 2010; Kushida 2011); and the role and construction of identity in interaction (Raevaara 2011; Schegloff 2017).

Beside those emerging three domains, Heritage and Drew saw continuing trends in the development of CA studies of specific types of interaction in institutionalized settings, workplace, and other settings which is compiled as the fourth domain (Clayman 2011; Kuroshima 2010). Between 2005 and 2010, there was a rapid growth in the number of publications in the three new fields of contemporary CA (shown as blue, green and orange color dots) and the continuing studies in institutionalized CA (shown as pink color dots) (Figure 1.2 from Drew and Heritage 2013).

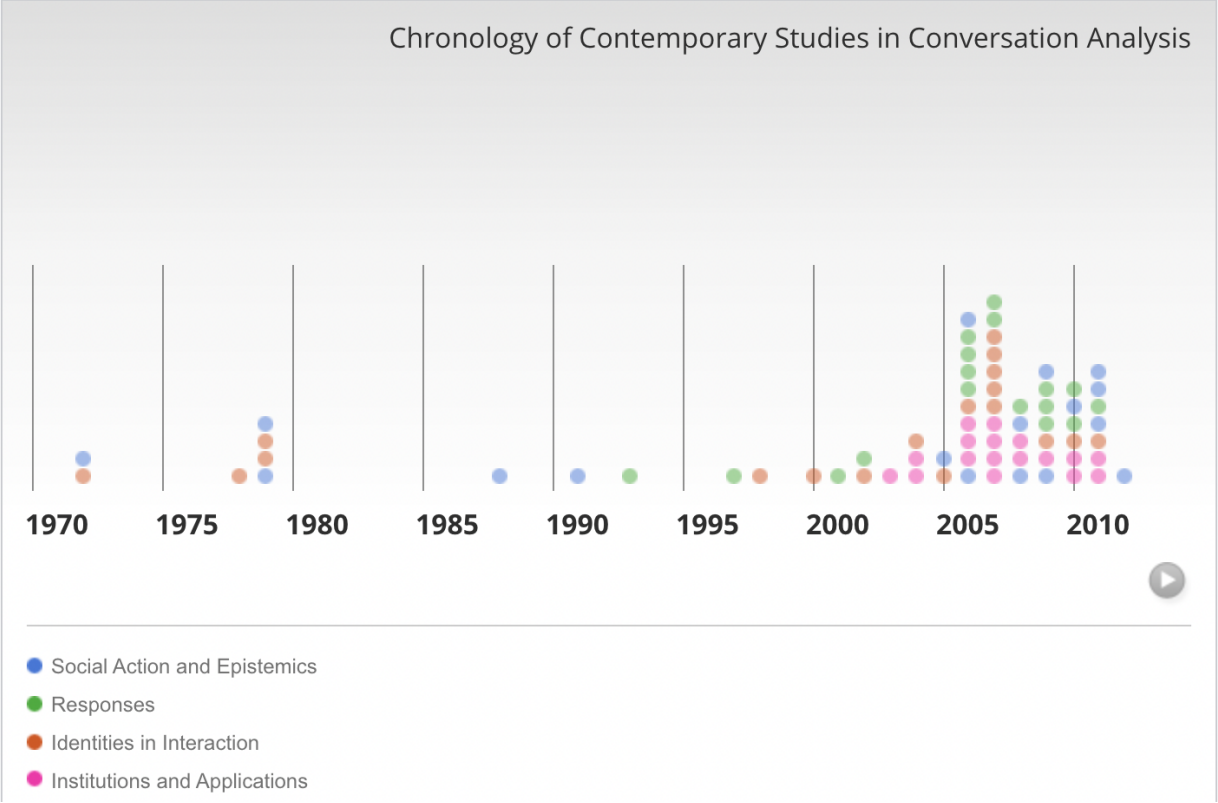


Figure 1.2 Chronology of Contemporary Studies in Conversation Analysis

More recent works include writing by social psychologists drawing on CA research and addressing non-specialist readers, such as *The Science of Conversation* (2018) by Elizabeth Stokoe. This book is largely aimed at helping readers have more effective conversations, and it

covers a range of topics including gender differences, while also translating many of the insights of CA researchers for a broad readership. Another such example is David Crystal's *Let's Talk : How English Conversation Works* (2020). While many topics in these kinds of books follow the classic analytic themes of CA such as turn-taking, they also include more recent currents of CA such as online conversations and misunderstanding that can happen in conversations due to cultural differences.

Selected themes of Conversation Analysis

In developing my own techniques of how to apply conversational models in music composition, I have chosen to use primarily two widely-cited texts, *Conversational Organization: Interaction between Speakers and Hearers* (1981) by Charles Goodwin, and *A Simplest Systematics for the Organization of Turn-Taking for Conversation* (1974) by Sacks, Shegloff and Jefferson, along with a more recently established study focused on classification overlap in conversation, *Nine Types of Human Interruptions in Conversations* (2020) by John Kane.

Turn-taking organization for conversation by Sacks, Schegloff and Jefferson

Sacks et al. propose a model for the turn-taking organization for conversation, referred to as a "speech exchange system," and examine its compatibility with a list of observable facts about conversation. They establish rule sets for the organization of turn-taking, based on two assumptions: 1) one party talks at a time despite the change of speakers, duration of each turn, and ordering of turns; and 2) transitions are finely coordinated using the techniques for allocating

turns. Below are the basic rules that govern the turn-taking construction, providing for the allocation of a next turn to one party, and coordinating transfer so as to minimize gap and overlap.

> Rule 1. For any turn, at the initial transition place of a turn-constructional unit:
> a) A current speaker selects next. The party selected has the right and is obliged to take the next turn to speak; no others have such rights or obligations, and transfer occurs at that place.
> b) If the 'current speaker selects next' technique is not involved, then self-selection for next speakership may, but need not, be instituted; first starter acquires rights to a turn, and transfer occurs at that place.
> c) If the 'current speaker selects next' technique is not involved, then the current speaker may, but need not continue unless another self-selects.
>
> Rule 2. If, at the initial transition-relevance place of an initial turn-constructional unit, neither 1a or 1b has operated, and, following the provision of 1c, the current speaker decides to continue, then the rule-set 1a-c re-applies at the next transition-relevance place. (Sacks, Schegloff and Jefferson, 704)

Stages of engagement and disengagement by Goodwin

In *Conversational Organization: Interaction between Speakers and Hearers* (1981), Goodwin introduces his interpretations and understanding of units of talk, i.e. "utterance." He uses the term "utterance" to refer to the stream of speech actually produced by a speaker in conversation, and the word "sentence" as well as related terms such as "phrase" and "clause" to refer to abstract entities capable of describing distributional relationships within and between utterances. In defining "utterance" as the actual stream of speech, he also means to include the entire vocal production of the speaker - that is, not only those sounds which could be placed in correspondence with elements of sentences, but also phenomena such as mid-word plosives, in-breathe, laughter, crying, "uh's", and pauses. While he distinguishes those types of utterances such as sentence and ones in the phonemic clause or breath group, he admits the possibility of an

utterance containing several sentences as well as possibility of a sentence being constructed through several utterances (Goodwin, 7). In other words, in his definition, an utterance can consist of vocal productions of the speaker in various sizes as little as "uh" and as big as groups of several sentences.

Goodwin analyzes organizations of engagement and disengagement in conversation through investigations of several cases. Participants utilize both their bodies and a variety of vocal phenomena to show each other the type of attention they are giving to the events of the moment, and reciprocally, the type of orientation they expect from others. In each case, co-presence of participants is sustained while the form through which they express to one another is not. How participants move from one of the alternatives open to them to another, and the consequences that such displays have for the organization of their talk can be observed in two contrasting cases, mutual engagement and mutual disengagement (Goodwin, 95).

Figure 1.3 shows displays of mutual engagement (Figure 3.1) and disengagement (Figure 3.2) (Goodwin, 97). Mutual engagement takes place when two participants are co-present and are engaged with each other. A gazes towards B, thus showing that B is being publicly observed and that A herself is positioned to take account of what B is doing. Mutual disengagement is when two participants are co-present but not engaged. A is gazing away from B. B is therefore not being publicly scrutinized by A, and A is not observably positioned to perceive all of the actions B might perform and incorporate them into her own actions.

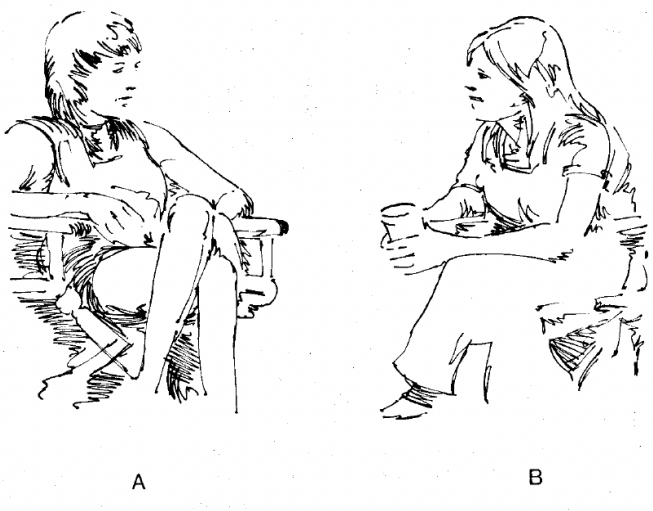


FIGURE 3.1

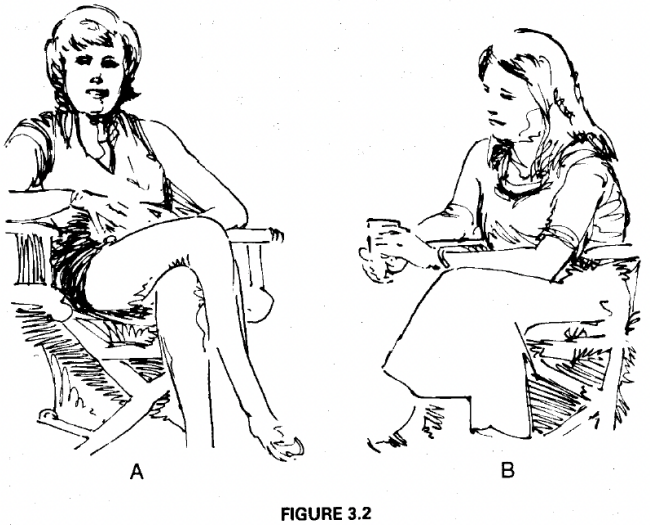


FIGURE 3.2

Figure 1.3 Displays of mutual engagement and disengagement

What can be noted here is that a display of a disengagement treats someone who is physically present as being, in a certain sense, not relevantly present, that is, not the subject of observation or a locus for joint, collaborative activity. Displays of engagement thus permit the alternation between presence and non-presence to be reestablished within a domain bounded by physical co-presence, and to become a relevant feature of activities occurring here.

Another point to note is that displays of disengagement show that, despite their displayed lack of orientation toward each other, the participants are in fact monitoring each other's actions quite closely. Their lower bodies remain oriented toward each other, even though the upper parts of their body, especially their gaze, are directed away from each other (Goodwin, 98). In addition to monitoring, despite the lack of orientation toward each other, participants are engaging in ongoing analysis of those actions. Hence, during periods of disengagement, the participants are situated in the same space, being engaged in the talk, having their lower bodies oriented towards each other while upper bodies are turned away. These positions and actions allow them to monitor and notice any changes made by another party. Goodwin states that regardless of displaying lack of orientation towards each other, "such displays of non-collaboration are in fact organized interactively and collaboratively sustained by the careful, systematic work of participants who maintain an ongoing monitoring of each other and an orientation toward the possibility of relevant changes in their mutual participation status" (Goodwin, 101). Possible behaviors observed upon entering disengagement include the following (Goodwin, 102-113):

1) Operations by recipient during talk: Nodding, for example, enables a recipient to display not simply hearership, but some aspect of his understanding of the talk then being produced. It might inform the speaker that the recipient's understanding of the talk is inadequate, then the speaker may modify the talk to obtain a more appropriate understanding. It only shows the talk has been attended to, and dealt with in a relevant fashion by its recipient. It does not deal with the issue of whether the space after the talk is or is not to be treated as a place where further talk is relevant.

2) Visible withdrawal from talk by speaker: This occurs, for instance, when a speaker withdraws her gaze and becomes silent, and changes the position of her cigarette. Though the speaker both withdraws her gaze and becomes silent, she is still actively involved in producing her talk but with less than full engagement. Recipients do not attend to the gaze withdrawal as an isolated event, but rather analyze it with reference to other activities the speaker is performing at that moment.
3) Recipient withdrawal: An example of this behavior is when the recipient begins to nod and withdraws her gaze while the speaker talks; the withdrawal is occupied by talk-relevant activities. Nods during withdrawal are not performed in the same way as those done during full-engagement; they are slower and more subtle. Though the recipient's nods continue to perform actions relevant to the talk, these actions are performed in a way that is sensitive to the changes in engagement states that are occurring.
4) Activity-occupied withdrawal: In this behavior, the recipient overlays talk-relevant acts with simultaneous moves away from talk. The speaker continues to talk while withdrawing from her recipient and positioning herself for entry into a different activity. There are two consequences of organizing a withdrawal in this fashion. First, the boundary between full-engagement and mutual disengagement is not structured as a sharp break. Participants are given a space where they can reorganize their bodies and actions to show the change and display their understanding of the talk. Second, the act of dis-attending each other never emerges for either party as a noticeable, recognizable activity in its own right. When each participant finishes her talk-relevant activities, she finds that she and her co-participant are no longer in orientation toward each other, that

state of affairs having been systematically achieved but never made visible as an explicit act of disaffiliation.

5) Refusing to withdraw: When speaker A is about to complete his utterance, the recipient B had started engaging in another activity, i.e. drinking a beer, but because A continued to gaze at B, B decided to put aside another activity, and saw A's continued gaze as proposing further talk from him was relevant.
6) Maintaining Availability: When A ends her talk without engaging in another activity, the recipient B withdraws her gaze from A, but rather than immediately moving to a state of full disengagement, B maintains her body position toward A, until B finds out that A doesn't produce further talk after a while. The behavior of B displays her acknowledgement of the possibilities that A may continue, but need not do so, as well as her availability to become the recipient should A continue.
7) Matching displays: Displays of eyebrow flash demonstrate attentiveness to the speaker and also that the recipient has in some way dealt with the particulars of the talk of the moment. It also allows her to perform explicit withdrawal by dropping the brows. When both the speaker and recipient perform eyebrow flash, they demonstrate a display of congruent understanding that not providing further talk may be relevant. This explains why such actions are often found at places of transition from talk to disengagement.
8) Talk with disengagement: In this behavior, the speaker makes her talk available but does not propose that the recipient should explicitly demonstrate that it is being heard. By doing so, the speaker also avoids displaying that she is awaiting a next utterance to it without indicating that subsequent talk would be inappropriate.

Studies around overlaps by Jefferson, Kendrick and Kane

The topic of overlap in conversation has been one of the focal points in CA, often included in the category of Turn-taking and Repair. One of the first analyses of overlap is Jefferson's *Notes on 'Latency' in Overlap Onset* (1986). Jefferson saw overlap as a byproduct of two activities: "1) A recipient reasonably, warrentedly treats some current utterance as complete, 'transition ready', and starts to talk, while 2) the current speaker, perfectly within his rights, keeps going." By locating a series of 'positions' across the transition space in conversation, Jefferson demonstrated cases of an instance of clean transition, and then instances of overlap (Jefferson 1986, 154). In this article, Jefferson breaks down overlap into five categories depending on the positions of overlap in the conversation or the intention of the participants (Jefferson, 154-175):

1) Possible completion onset: The recipient has a particularly good warrant to treat an utterance as completed or transition ready. Here, the overlap can be minimal and unproblematic, as in the following two cases (Jefferson, 154):
    a) Terminal onset: Two people starting up simultaneously.
    b) Last item onset: A case is resolved by the current speaker dropping out.
2) Interjacent onset: A recipient would start up just after the current speaker had produced a clear indication of going on, following a possible completion (Jefferson, 158).
3) Unmarked next position: A recipient/next speaker produces his talk in such a way that it occurs with neither haste nor delay. It is not pushed up into or latched immediately onto the prior utterance, but permits just a bit of space between the end of the prior and the start of the next (Jefferson, 162).

4) Consensus or reaction: When completion of an utterance is acknowledged with consensus, ambiguity or disagreement (Jefferson, 167)
5) What is completion: Various treatments of silence between utterances (Jefferson, 175)

For each type of overlap, more than ten examples of conversation are analyzed with transcriptions, each exhibiting slightly different situations which can be problematic. For instance, the basic assumption of the first category, "possible completion onset," is a clean transition with little or no overlap. However, examples of cases here show several instances of overlap, some of which are short response utterances such as “yes” or “ah hah.” In other cases, “huh” at the end of the speaker’s utterance overlaps with “I don’t know” of the recipient, or both the speaker and recipient start talking at the same time. Since I see these to be demonstrating different cases, it was not clear on what criteria they could be assumed unproblematic. In addition, there may be duplicates or instances that closely resemble each other in more than one category. Short utterances such as “ah”, “oh” or “mm” are discussed in "unmasked next position" as a form of acknowledgement which occurs in the midst of someone’s talk (Jefferson, 163), while cases 1.7, 1.8 and 1.9 also deal with utterances such as “yes” or “oh” during the turn-taking process to show consensus (Jefferson, 156-157). While Jefferson’s detailed evaluation from transcriptions of conversation is highly valuable and demonstrates the delicate and intricate nature of overlaps in conversation, these analyses could benefit from further refinement.

In *The Timing and Construction of Preference: A Quantitative Study* (2015), Kendrick and Torreira conduct highly contextualized research of conversation concerning the timing and construction of preferred actions (e.g. acceptances) and dispreferred responding actions (e.g.

rejections). Though this research focuses on examining the duration of the gap (silence) between turns, instead of actual overlap, I consider this to be a related topic, as it concerns the degree of preference or dispreference depending on the timing. One clear prediction that the authors derive from prior research is that the duration of the gap between the end of a first part and the beginning of a second part should be systematically greater for dispreferred actions than for preferred actions (Kendrick and Torreira 2015, 260). The first conclusion the two authors draw from their results is that the size of the gap, in other words, the timing, is a variable signal which impacts the frequency of receiving preferred response actions or dispreferred response actions. In particular, a small departure from a normal turn transition (i.e., a gap of between 300 and 700 ms) may alert the recipient that the most frequent response type, a normal acceptance, is less likely. On the other hand, with a relatively large departure (i.e. a gap of 700 to 800 ms or more), the recipient may expect dispreferred response actions (rejection) to be more imminent (Kendrick and Torreira, 287). The authors' second conclusion is that the timing of response can be understood as a "turn-constructional feature" which is "the virtual component of a preferred or dispreferred turn format".

Offering an alternative model, John Kane's "*Nine Types of Human Interruptions in Conversations*" (2020) examines human interruption in conversation, breaking it down into nine different types. This classification focuses specifically on overlap that occurs during turn-taking. Kane proposes nine distinct classes of overlap:

1) Smooth transition: Both speakers are able to finish their turn, without interference. There may be a small overlap but starting early did not affect the conversation.

2) Collaborative overlap: One person starts talking before the other person finishes and because of this, the person who was speaking first has to alter what they were going to say or even stop speaking prematurely.

3) Competitive overlap: One person starts speaking before the other finishes and the interruption is not welcome. This interruption can sound competitive, meaning that the first person is reluctant to stop talking because, for example, they disagree or the conversation changes course.

4) Collaborative pause overlap: A person may pause during a conversation, and wants to continue but before they can do so, a second person starts speaking. Here, the original speaker lets the other person continue, making the interaction sound collaborative or neutral.

5) Competitive pause overlap: Similar to "collaborative pause overlap," a person might pause in conversation but before they continue, the other person starts talking. In this case, the original speaker may or may not attempt to interrupt themselves.

6) Speaking butting-in: One person is talking and another begins with the intent of talking over the conversation, however, the original person continues to talk regardless.

7) Pause butting-in: In this instance a person pauses during a conversation, and the other begins to talk, but the original speaker continues on with their thoughts. The second person stops talking and allows the other to continue.

8) Backchanneling: Similar to "Speaking butting-in," one person chimes into the conversation, but the intent is not to take over the conversation. This might include them

saying anything, as long as it is said in a way that allows the original speaker to continue (examples include “yeah”, “uh-huh”, “oh really?”).

9) Accidental overlap: After a prolonged silence both people might start talking at the same time, but eventually one of them gives up and one carries on the conversation.

I find Kane’s analysis highly accommodating and complementary to the theories by Sacks et al. and Goodwin in that it brings a different approach to turn-taking, yet provides practical and accurate observation of conversation. Whereas Sacks et al. deal with rules of turn-taking, and Goodwin focuses on stages of engagement and disengagement in conversation, Kane highlights the complexity of human interruptions, categorizing overlaps during the course of conversation based on the dynamics of both the speaker and recipient's intentions, actions and responses.

Though the theories and classifications presented by Sacks et al., Goodwin and Kane seem sufficient to cover most situations in conversation, one additional case that must be acknowledged and is relevant to musical composition is when multiple speakers talk. In group conversation or discussion, it is inevitable to encounter a situation in which a participant begins to talk over the speaker which triggers another participant to join, and so on. Or multiple participants could start to make utterances simultaneously or randomly creating a situation in which multiple participants turn into speakers. This could be classified in a number of ways: 1) multiple participants begin making utterances in response to one speaker, hence everyone is engaged in the same conversation; 2) multiple participants make utterances at their own will, overlapping with utterances of other participants without the presence of a main speaker; and 3)

in relatively large group settings or on the street where random people pass by, there could be more than one conversation occurring at the same time, and there may be people who are not engaged in any of the conversations or are about to disengage. While I have not yet found any scholarly works that focus on conversations organized in this way, these conversational scenarios are commonly observed in music making. For example, contrapuntal textures in European classical music involve instruments or voices simultaneously performing multiple melodic lines considered to be of equal importance; likewise, many common textures in jazz and improvised music highlight conversational modes that depend on similar techniques of polyvocality. In the next chapter, I turn to questions of how Conversation Analysis and conversation more generally can be used to understand compositional and improvisational techniques.

Chapter 2: Analysis of musical composition and improvisation through the lens of conversation

Collective improvisation and overlapping conversations

Studies of conversation generally focus on turn-taking with one speaker at a time and events of overlap are treated as somewhat special, in that they require particular actions and responses. Overlaps in music when used as a reflection of conversation, on the other hand, can be treated in numerous ways to produce a variety of outcomes. For instance, collective improvisation in free jazz can reflect a particular dynamic of conversation that is common in music, but considered problematic in conversation. In this section, by investigating the nature and the construction of collective improvisation, I will research how such situations of overlap in improvisation are treated by some of the composers in the 1960's by looking into approaches by Charles Mingus and Ornette Coleman, and in what way they reflect aspects of conversation.

The concept of collective improvisation has been explored by jazz musicians; as Gunther Schuller notes, "the improvisation of many lines at the same time is a typically African concept, and is perpetuated in most forms of early jazz, a music marked above all by 'collective improvisation'" (Schuller 1968, 57). It gained attention as it emerged as part of the evolution of free jazz at the beginning of the 1960's (Angelino 2020, 50). Collective improvisation is, simply put, simultaneous improvisation by several musicians in which all performers' expressions are considered of roughly equal importance, as opposed to more soloist-accompaniment approaches to musical texture, as in conventional forms of jazz in which one player improvises over chord

changes while accompanied by the rhythm section. The original idea of collective improvisation, "group improvisation" in Ornette Coleman's words, dates back to New Orleans' early bands where it played an important role, but this changed in the music of big bands during the swing era (Coleman 1959).

Collective intentionality in collective improvisation

Approaching 1960s free jazz from a philosophy perspective, Angelino relates collective improvisation to current discussions around "collective intentionality," for several reasons. First, collective improvisation is identified as "complexly interwoven sets of collective intentions" (Hagberg 2016, 484) which implies that it is not merely a sum total of its parts but is interweaving intentions of the performers through spontaneous interactions. Second, the players have the liberty to contribute what they feel in the music at any given moment, but at the same time, they perform collaborative actions. In this regard, the basic construct of collective improvisation lies in a different dimension compared to the application of improvisation in the traditional format of performance in jazz, in which the format of performance from the head to solo is already built in and performers play their roles accordingly regardless of their intentions. In Ornette Coleman's words, in the standard format of improvisation "the individual is either swallowed up in a group situation, or else he is out front soloing with none of the other horns doing anything but calmly awaiting their turn for their solos" (Coleman 1959).

Collective improvisation by Charles Mingus and Ornette Coleman

Charles Mingus, who built a vast body of work featuring collective improvisation from the 1960's, established specific concepts about collective improvisation that he required his players to follow. Berliner refers to:

> "...a Charles Mingus composition that depicts the chronological history of jazz, requiring band members to remain faithful to the style of each arranged episode from ragtime to the avant-garde. Representing another approach, groups may combine the performance practices of bebop and free jazz by altering improvisation inside and outside a piece's structure. Soloists may, for example, periodically perform free or atonal passages against the rhythm section's conventional harmonic accompaniment, a dramatic effect that maximizes the tension between the parts before soloists return to the piece's structure. The rhythm section may follow the soloist's lead in temporarily abandoning, then returning to, a piece's harmony" (Berliner, 338-339).

These imply that Mingus as a composer provided a framework such as styles of music or tonalities for the players to choose from and to improvise. This allows the players to express themselves within the style or tonality of their choice while also increasing the sense of intentionality within the players as they are made aware of their responsibilities to contribute to the group's collective goal. I see this as a scene of group discussion in which all participants are granted the freedom to express themselves, agree with someone or disagree. Intentionally choosing to impose a different style than the rest of the band may be considered a gesture of overlap. Similarly, playing free or atonal passages against the rhythm section or some other players playing in conventional harmonic context may demonstrate a gesture of disagreement. Regardless of the intentions, these requirements to consciously choose the style or tonality to opt in or not is the key to keep the collective improvisation organized.

Ornette Coleman, on the other hand, approaches collective improvisation in a contrasting way using his system of “motivic chain associations,” which is a methodology to generate phrases by linking fragments of a provided phrase to another which I will discuss further below (Jost 1994, 50). The term “free jazz” arose in part due to the release of Coleman’s historic album *Free Jazz: Collective Improvisation* (1960) for double quartets (Jost, 59). To record *Free Jazz,* two full quartets gathered in the studio of Atlantic Records. The musicians joining Coleman were Don Cherry, Charlie Haden, and drummers Billy Higgins and Ed Blackwell, plus Eric Dolphy on bass clarinet, trumpeter Freddie Hubbard and bassist Scott la Faro, who carried Coleman’s conception of “free group improvisation” to a “totality such as had never been heard before in Jazz" (Jost, 59). The music featured collective improvisation by the two independent quartets, each containing two wind instruments and a rhythm section of bass and drums. Having two bands, each playing their improvised melodic, harmonic and rhythmic contents in their own tempo and not always in alignment with what the other band is doing, does create a sound that most listeners would consider chaotic. However, the outcome is not constantly additive. Rather the music has dynamic textures that there are places where some players lay out, letting only the rhythm section or the horns in one band play. These behaviors not only create various dynamics in texture and surprises in the performance but also illustrate interplay between individual players or between sections of the groups as if they are talking, listening and responding to each other.

Analysis of *Folk Forms No.1 and* Coleman's *Free Jazz, A Collective Improvisation*

Here I will examine Mingus's *Folk Forms No.1 and* Coleman's *Free Jazz, A Collective Improvisation,* drawing on the work of both Jost and Angelino. As Jost points out, despite the different musical results of these two composers' conceptions, there is one tendency concerning the evolution of free jazz that is observed in both works: "the move from individual monologuing soloists towards a kind of collective conversation" (Jost, 61). In that regard, despite his sometimes critical comments about his contemporaries who were associated with more "avant-garde" forms of jazz, Mingus is recognized as an influential figure in the development of free jazz, and his *Folk Forms No. 1* is considered as one of the most important milestones in the evolution of new approaches to collective free improvisation in which the traditional roles of rhythm section and soloists are set aside and altered in a way that is not seen even in the early music of Coleman (Jost, 42-43).

Drawing on the example of *Folk Forms No.1,* I will examine the conversational aspects and how individual players maintain the sense of collectiveness/collaboration. During its collective improvisation, there is almost no steady beat heard and hardly any continuous harmonic structure. The music also lacks a clear sense of fixed accompanists/soloists roles. Hence the mood can be generally hectic or nervous, but is never chaotic (Jost, 43). What is worth noting is that while being "on a par with all the others," individual players can adapt themselves to the rest of the band, forming "conciseness" which stems not so much from any convention or pre-planning, instead arising "from an intuitive understanding between musicians." (Jost, 43) As explained by Dannie Richimond, the drummer in Mingus's group, players in Mingus group would "feel each other out" as opposed to making agreements or planning regarding the

interactions during the performance beforehand (Jost, 38). By performing in such a way, the nature of the performance becomes "autonomous but *not independent* but; as in a conversation, there are pauses, and the explanations of one player are corroborated by another" (Jost, 43). Jost points out how players corroborate each other during the performance in *Folk Forms No. 1* (Figure 2.1). The accents on right before beat four set up by Mingus (double bass) in bars 1 and 3 is taken over by Eric Dolphy (alto sax.) in bad 5 and Ted Curson (trumpet) in bar 9. These gestures imply the readiness and willingness of the players to support other members of the group while each freely plays and expresses what they want. In other words, all the players are intensely monitoring each other's behaviors by simply being in the same space focusing on the moment without needing instructions as to who plays what and when. This itself represents a group conversation that is unscripted yet self-organized. In addition, the gesture of corroborating other players can be seen as displays of attentiveness or agreement by making short utterances in conversation as backchanneling not intended to interrupt but to support.

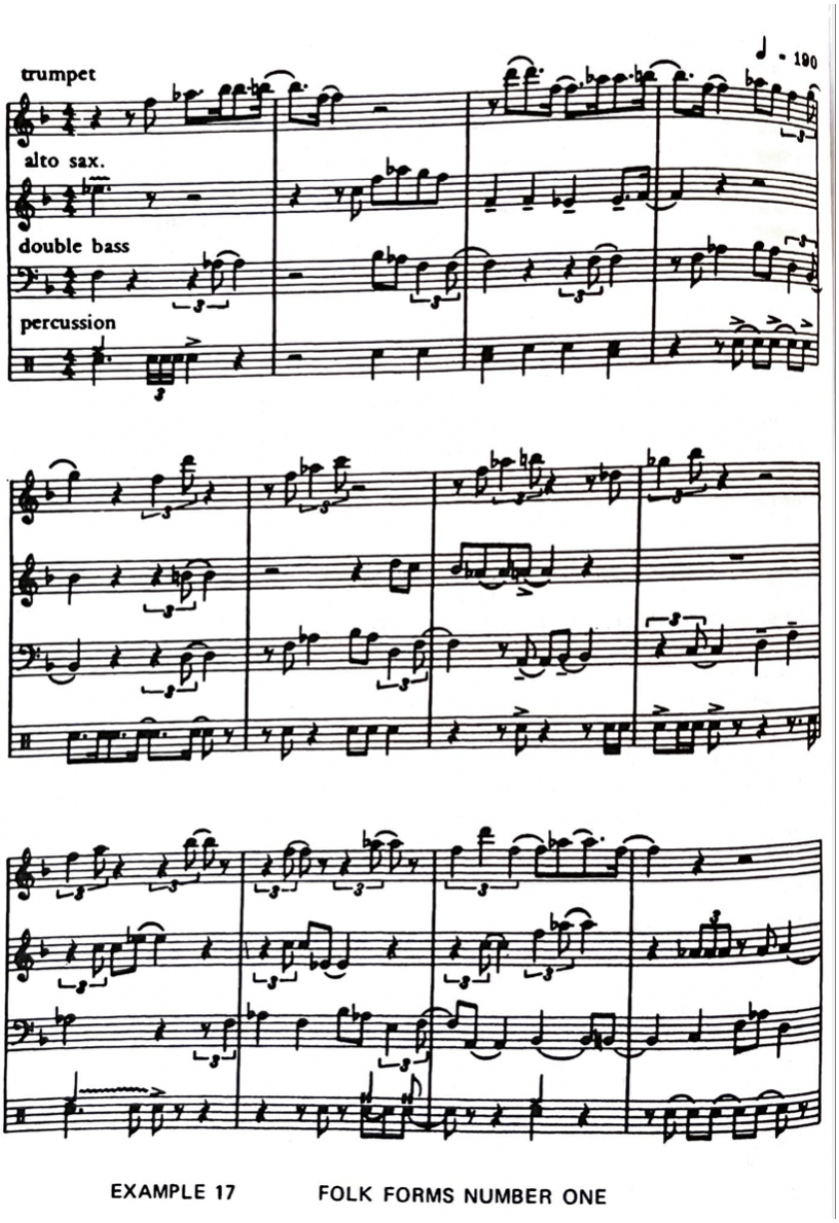


Figure 2.1 Transcription of *Folk Forms No.1* by C. Mingus

Another approach seen in Mingus's collective improvisation that implies conversational elements is the spontaneously changing groupings. As opposed to changing from one soloist to the next, in the case of Mingus's group, a highly variable sound is achieved by intuitively and spontaneously changing groupings, particularly in large formations but also as a quartet—

creating "pieces in which no two choruses are alike" (Jost, 40). For example in *Fork Forms No. 1,* duets—trumpet and alto (unaccompanied), trumpet and bass, or alto and drums—are formed between the ensemble passages, thus constantly creating timbral and dynamic gradations and shifts that are always different (Jost, 40). In conversation, this would be interpreted as a situation where speakers and hearers spontaneously withdraw from the state of engagement to disengage and to enter into another state of engagement with a different participant. As speakers and hearers enter or withdraw from the states they are in, they perform a series of gestures that can be eluded in both implicit or explicit ways and are constantly changing, which in music would create unique combinations and phrases of change each time.

Coleman's approach to collective improvisation is significantly distinct from that of Mingus in which cohesiveness and direction are consciously implemented in the music by the composer as a framework (Angelino, 53). In his liner notes for *The Change of the Century* (recorded in October 1959) Coleman describes:

> "When our group plays, before we start out to play, we do not have any idea what the end result will be. Each player is free to contribute what he feels in the music at any given moment....
>
> I don't tell the members of my group what to do. I want them to play what they hear in the piece for themselves. I let everyone express himself just as he wants to. The musicians have complete freedom, and so, of course, our final results depend entirely on the musicianship, emotional make-up and taste of the individual member. Ours is at all times a group effect and it is only because we have rapport we do that our music takes on the shape that it does. A strong personality with a star-complex would take away from the effectiveness of our group, no matter how brilliantly he played."

As stated above, Coleman does not provide stylistic requirements for improvisation. Instead, by granting complete freedom to the performers, Coleman is, in a way, letting go of his

control as a composer, solely relying on the rapporto between the players to allow the musicianship of the performers to be the creators. Again, I see these activities of performers as a form of group discussion in which each individual is saying what they have to say to contribute to the group as a collective. However, in the case of Coleman's collective improvisation, it is performed with much less constraint.

On the other hand, Angelino raises a question about the intentionality behind the outcome of: "How do players come to act, feel, think and perform *qua* members of a team?" (Angelino, 53). While Angelino acknowledges recent decades of studies on how "collective intentionality" governs individuals' behaviors, including work by researchers such as Michael Bratman, Raimo Tuomela, John Searle and Margaret Gilbert, she draws a clear distinction between these theories and the actual events during musical performances, implying that there are characteristics in the case of collective improvisation in music that do not always align with the theories (Angelino, 53). Instead, Angelino makes the following remarks: (Angelino, 55)

- First, players begin without a pre-conceived notion as to what kind of effect or goal they want to achieve. Consequently, what happens in the music is not planned in advance, but arises from improvisation (e.g.: from spontaneous interactions within the group), and is molded to a large extent by the emotions of people who play it. ...
- Second, each player is "free to contribute what he feels in the music at any given moment." not only in solos, which will be individual in any case but, more importantly, in collective playing.
- Third, players are not committed to a peaceful dialogue. Rather, they are also adversaries who challenge each other and challenge themselves to listen and take note of what they are saying musically.
- In spite of this the fundamental freedom and in spite of this multi-directional, open structure, an inner coherence is created through a free association of melodic lines, in such a way that the fusion of sonic forms in both Mingus' and Coleman's work emerge as surprisingly spontaneous and startlingly logical at the same time. In the words of Coleman, "there is a continuity of expression, certain continually evolving strands of thought that link" all parts together.

While these statements of Angelino refer to the musical context of collective improvisation, they constitute as fundamental principles to hold constructive and open group discussions. The first point implies having no planning and assumptions prior to the discussion and allowing emotions of the people to develop the conversation. The second point lays the basic assumption for the group that each participant is free to talk at any given moment while being aware of the collective goals. The third point suggests the possibility of disagreements, interruptions and overlaps that may happen during the conversation. If that happens, the participants are expected to listen and respond appropriately. On the other hand, the fourth point seems unique to the musical context, and cannot be converted into a context of conversation directly. However, such open structure and inner coherence in conversation can be created through not a free association to melodic lines, but participants carefully responding to each other's utterances to build larger context.

Mingus would also demonstrate ideas of the music at the piano as a framework of tonality and some ideas about rhythmic values of the notes for creating coherence in collective improvisation instead of providing harmonic or melodic structures. In a way, this sets up a general assumption or overall tone of the talk for the group. As Jost explains, "Mingus brings his musicians directly into the compositorial process; he does not compose *for them,* but *with them.* He does not hand them a definitive score, but sketches for them—usually at the piano—the musical and emotional framework of a piece." (Jost, 38) While providing such a framework, he leaves plenty of room for the players to make their own choices to contribute to the music in their own way. As Mingus puts it "As long as they start where I start and end where I end, the

musicians can change the composition if they feel like it. They add themselves, they add how they feel, while we're playing" (Quote by Honoff. Jost, 38).

As Angelino points out, consequently, such an approach in performance based on these agreements demands readiness to interact from the players; the absence of the convention of a pre-composed melodic and harmonic structure forces the members of a group to listen to each other with intensified concentration (Angelino, 56). Each player needs to foresee what might come ahead of them by paying close attention to other players' activities, and be prepared to adapt or make adjustments whenever needed. Again, this is an innately built-in feature in conversation as well as a requirement if participants want to hold the conversation since monitoring of each other occurs whether one is in the state of engagement or disengagement while being part of the conversation.

Coleman, on the other hand, utilizes his strategy "motivic chain association" to improvise while maintaining overall sonic coherence and avoiding the risk of falling in chaos, in which "one idea grows from another, is reformed, and leads to yet another new idea" (Jost, 53). In motivic chain association, linkage is created in the ending of a given phrase to a new phrase derived from it using the rhythmic and melodic pattern. Beside its innovative way of generating phrases, one important aspect of this motivic chain association is the collaborative interplay it naturally evokes between the players. As the linkage is created at the end of a phrase introduced by a player, it provides possible implications or clues to the next player, who then picks up one of the given implications to generate his phrase, which will then create another set of possible implications for the next player—causing a cyclic process that continues to evolve. Angelino calls this process "collaborative emergence" as it arises from interactions among the members of

the collective (Angelino, 57). Another important point regarding motivic chain associations is that they are not limited to occur linearly from phrases that follow one another immediately. They can take up ideas that are several links back in the chain, creating larger contexts in this way (Jost, 50). Hence, examples of motivic chain associations can be demonstrated in an indefinite number of ways. Not only do they not rely on harmonic relationships, they offer moments of surprises and conflicts by imposing motifs in a way that may run beyond one's expectations (Angelino, 58).

In a conversational context, motivic chain association can easily be interpreted as turn-taking in a dialogue or in a group discussion where there is one speaker making an utterance and the next speaker formulates a new utterance making an association to what was said by the previous speaker. To make a linkage in turn-taking, the next speaker can paraphrase the previous speaker's utterance at the opening of his turn before adding his own inputs or simply using some of the same words to create his utterance to continue through linking. Such turn-taking with linkages will be a way to ensure the conversation to develop with coherence while displaying attentiveness to the current speaker rather than letting it terminate or change direction at any given moment at participants' will.

Summary

I have focused on the principles and strategies of collective improvisation, to investigate how overlaps in music (simultaneous voices) are treated, by surveying works of the two pioneers in free jazz: Charles Mingus and Ornette Coleman. To summarize, collective improvisation can be approached in a number of ways to maintain cohesiveness and direction of the piece, by

providing a framework—such as styles or tonalities of the music—and by implementing the realtime composition technique "motive chain association". Within those frameworks, performers are granted freedom to express what they want to say and even disagree with each other or go against others while being aware of their tasks to contribute to the collaborative work. I propose two reasons why overlaps, even in the event of disagreement, can be reconciled in music. First, moments of simultaneous voices in different tonalities or styles adds temporary tensions, which will resolve eventually, and creates contrast and dynamics in music. Second, behind such gestures of disagreement or competitiveness in performance, all performers are aware of themselves being part of the collective and its purpose—to serve the music. Thus, any interplays between performers during the performance, whether collaborative or competitive, would evolve under the mutual assumption as a collective. On the other hand, in a context of conversation, overlaps are generally considered troublesome and need to be corrected before proceeding. Multiple speakers talking simultaneously can only add confusion or frustration, regardless of their intentions. Although I will refrain from investigating further as to why overlaps can work in music but not in actual conversations, as this is beyond the scope of this dissertation, my assumption falls into the very principle of turn allocation technique that "one speaker speaks at a time" (Sacks, Schegloff and Jefferson, 705). Overlaps do occur in conversation and are impossible to obviate completely. However, they are considered to disrupt the ordering of turn-taking. Therefore, they need 'special treatment'. As opposed to music, which can accommodate simultaneous voices, overlaps in conversation will need corrections because generally one can only comprehend the speech of one speaker at a time.

Composers creating scores for structuring conversational improvisation

Picture in Motion (2007) by Lisle Ellis

As an example of implementing collective improvisation in a score, I will examine *Picture in Motion* (2007) by Lisle Ellis, a score designed using his graphic notation system *Formings* (Figure 2.2). This system creates structured improvisation for the players, and at the same time, it challenges certain conventions of "free" improvisation. Instead of 5-line staff systems, a grid is used as a timeline to guide players when to begin, enter, overlap, exist and end. The graphic symbols in the timeline grid indicate what and when to play, when to not play and when scenarios such as solos, duos, trios, and other groupings take place. Beside what content to play and its context, those systems also function to assign individual or collective roles, contextualizing those roles within the piece. For instance, one graphic symbol can instruct a player individually, or as part of a grouping, to provide accompaniment to another player or group of players, or ask a player to engage with other players in call and response. The phenomena of silence is also addressed in the *Formings* including a group silence, pause, sudden halt or full stop.

This notation system functions beyond the conventional notation system by organizing the structure and defining orders, actions to be taken or not taken, roles and groupings, which Ellis says could be integrated with traditional notations. However, this system allows the players to interpret the instructions intuitively, accurately and efficiently using the codified graphic symbols in such a way that is not possible in the conventional notation system, which would only allow left to right timeline and would require much text explanation.

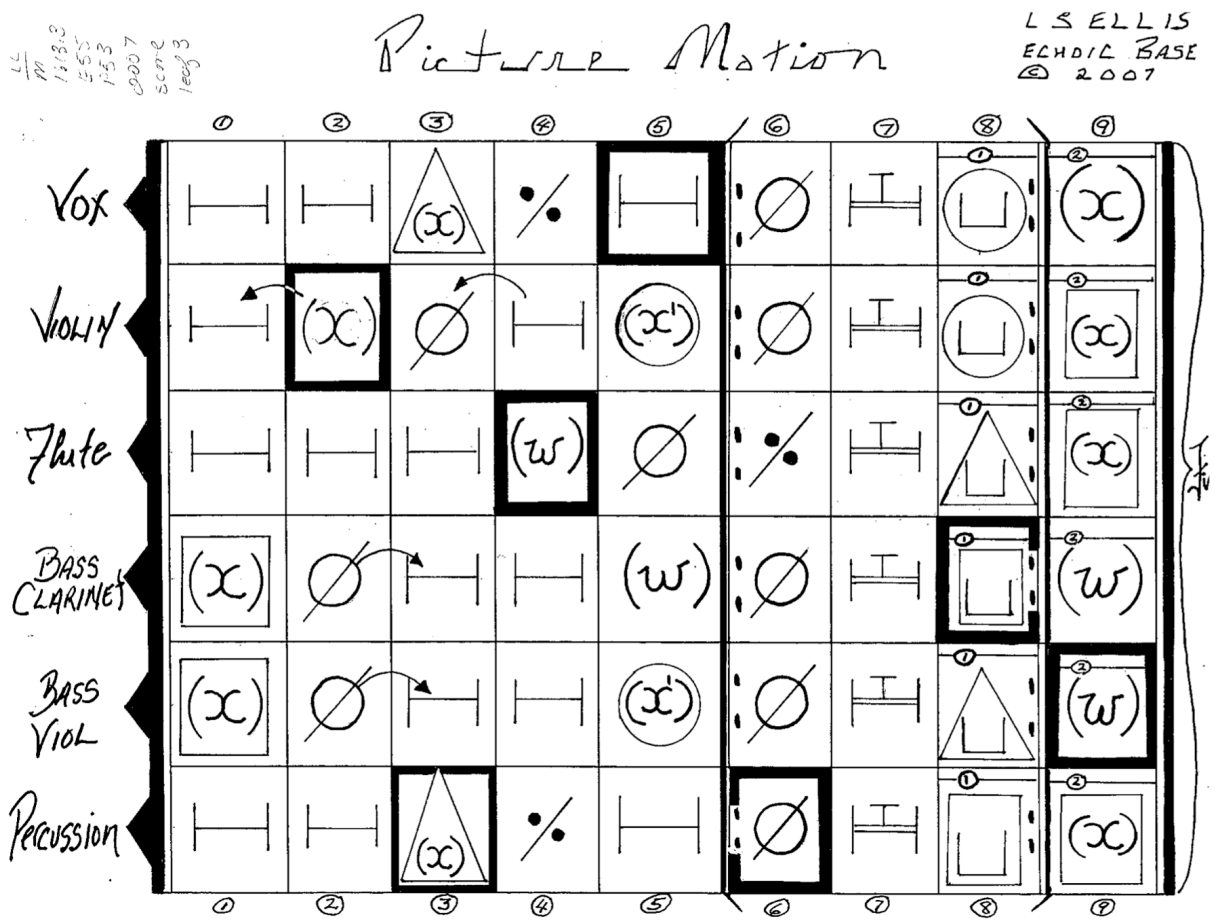


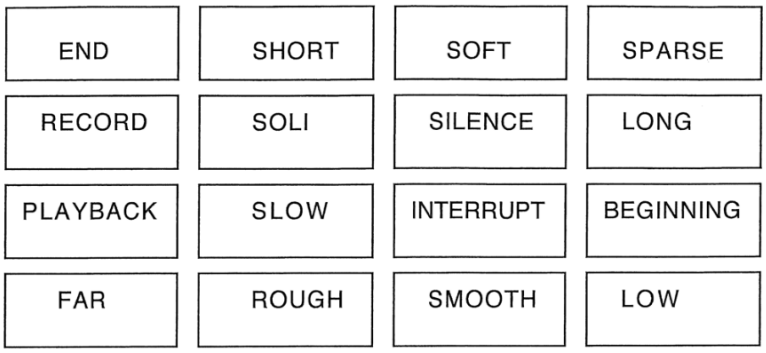
Figure 2.2 Score excerpt, *Picture Motion* by L. Ellis

Even though this grid works as a framework with certain instructions to the performers, the majority of the decisions in regard to performance still depend on the performers themselves. Ellis explains "remaining true to the spirit of improvisation, choices concerning pitch, rhythm, tempo, dynamics, and timbre are most often given over to the discretion of each player," allowing the performers to choose how they contribute to the co-creation. With regard to the conversational aspect, this framework is what forces the performers to communicate rigorously regardless of the assigned instructions for their parts—all players are asked to make choices to play/speak and to pay attention to what others are saying and react when needed. The next

actions are triggered by sonic cues indicated in the bold-boxed cells. When the symbol in the bold-boxed cell is played, that indicates the other players to move onto the next column. This requires performers to initiate actions that lead the group, collectively sharing responsibility for the pacing of the music, and as conversation, it demonstrates the participants paying close attention to each other communicating and collectively responding to the speaker's utterance.

Artificial Life 2007 by George Lewis

In *Artificial Life 2007*, George Lewis demonstrates a model of group improvisation based on situational composition for an ensemble of eight or more players. The way sonic environments are created in this piece as individual performers play their own parts without engaging with other performers directly but consequently blending through the sonic environments, is related to instances in *Group Talk No. 3*. In section A of *Group Talk No. 3*, performers create swells of sounds at their own pace without interacting with each others, creating multiple instances of sounds that continuously emerge to depict talking with disengagement and changing status of engagement. As a result, those sounds begin to merge and transform into another sound as the combinations of the sounds change. Although Lewis's piece did not originate from conversation, performance with intentional disengagement within the ensemble yields outcomes similar to my piece. *Artificial Life 2007* includes two different 1-page scores, each with text-based instructions. The score has a number of boxes with words that indicate specific procedures for structuring musical communication, facilitating the articulation and relationships to allow performers to create music in real time (Figure 2.3).



BEGINNING: Enter at the same time as another group, imitating them as closely as possible. Stop playing when the other group does.
END: Enter as soon as possible at the end of another group's playing, attempting to play what they just played as faithfully as possible, for the same duration as the original phrase.
FAR: Imitate the music being played now by the group farthest in physical distance from you. Stop playing when the other group does.
INTERRUPT: Play a very loud, raucous phrase for a maximum of 3 seconds, designed to interrupt another group's playing with a contrasting element.

LONG: Repeat a single short phrase for a relatively long time.
LOW: Play one extremely (relatively) long sound, pitched relatively low.
PLAYBACK: Repeat several times a phrase from another group that you previously memorized or recorded using the RECORD instruction.
RECORD: Record or memorize a phrase from another group and save it in your personal memory for future use with the PLAYBACK instruction.

Figure 2.3 Score excerpt, *Artificial Life 2007* by G. Lewis

The pacing and the order of playing are determined by the performers. What is most unique about this piece is that despite being an ensemble work, the music asks the performers to solely follow the instructions focusing on their parts and avoiding interactions with others unless instructed otherwise. From a conversation aspect, this implies intentionally-encouraged continuous speaking butting-in as performers are expected to keep ‘talking’ without engaging with other performers. Nevertheless, the performers are expected to be aware of their sonic environment since as Lewis describes in the notes, this piece is a "situational-form musical composition"; although performers do not interact with each other directly, the sound environments created by each player begin to emerge and create different combinations. This is

related to the changing states of engagement and disengagement theorized by Goodwin (Goodwin 1981). Similar to participants who continue to talk despite not being fully engaged in the conversation, but are in fact closely monitoring each other (Figure 1.3), the performers in this piece play their own parts at their own pace without interacting with each other, while being fully aware of the sonic environment that they are in. By moving from a situation to another situation, the performers experience different stages between full engagement and full disengagement. As this occurs within individual performers, new sonic environments will arise continuously creating different combinations of situations as a musical outcome. In this sense, the co-presence of different situations and environments that keep emerging and changing without being engaged or distracted, illustrate what from a Conversational Analysis standpoint could be described as a type of an overlap and talk with disengagement.

With regard to the goal and attitude of the performance, Lewis emphasizes the importance of following the instructions, rather than excellence in performance, disregarding one's familiarity with improvisation or background in musical training.

> "Artificial Life 2007 is open to performers from any musical tradition including those that do not regularly include improvisative modes of performance. Extensive prior experience with improvisation, while welcome, does not necessarily confer an advantage in performing this deceptively complex work; the best results will be achieved by following the instructions provided with care. The instructions serve as a kind of toolbox for producing a range of sounds and forms that will far exceed what the composer would imagine. For that reason, there is no canonically correct way for the piece to sound. Rather than assuming that recorded or live versions of the work might constitute a model for how the work should sound, performers should assume the freedom to create what they want to hear from a combination of the tools provided and their own creative and cultural standpoints.
>
> Even so, as with all improvisations, including our everyday-life human efforts, the success of a given performance of this work will be less a question of

> individual freedom than of the assumption of personal responsibility for the sonic environment" (Lewis, 2007).

In the last statement, alluding to improvisation as a part of everyday life activities, Lewis thus metaphorically treats this situational work as the representation of such an activity, one in which individuals collectively create an "environment." This relates to my idea that I am implementing in my own work, which is to manifest everyday conversation as a form of musical conversation. Just as the performance of this piece could yield outcomes that will exceed the composer's imagination, our everyday conversations with contributions from participants can create results that were not planned or expected, a dynamic that I seek to explore in my conversation-based compositional models.

Multiplexor (2006) by Victor Adan

In the score of *Multiplexor*, composer Victor Adan uses graphic notation to express sounds of the computer part, which cannot be notated with the conventional notation, in combination with a more traditionally notated flute part. I, on the other hand, used graphics as gestural notation for the performer, coupled with a traditionally notated piano part, to describe how the pianist should interact with the computer to recreate a dialogue between the two.

This score comprises mainly two parts, a main score being the bottom few systems grouped together which resembles traditional notation, and a second score located above it which is a graphical synthesis of the musical parameters involved (Figure 2.4). The second score is not a substitute for the first but rather it is an aid or a complement for the main score. Hence the

performer is required to follow the main score with precision. One of the intended purposes is to serve as a kind of "aerial map," used as an aid in understanding the overall structure of the piece. The second purpose is to serve as an analogue, although very coarse and almost cartoonish, representation of the general sounds resulting from the execution of the analytical score (Sauer 2008, 13). The traditional notation-like score coupled with the graphic score provides more precise performance instructions while allowing the performer to interpret the work more intuitively. The main score (traditional notation-like score), offers general yet clear ideas about the choice of the pitch without indicating specific notes, and can be adapted to the range of the performer's instrument or her choice of the range. It also gives exact information about the rhythm and the activity level. Though the rhythmic value is not exactly notated, with the provided tempo and measured bars, the performer can intuitively grasp the duration of the notes and the intensity of the activity required in each measure.

In my motion sensor piece *Cricket Wind*, I have adapted two staves similar to the score of *Multiplexor*: one for the piano using traditional notation and the other for computer and hand gestures with graphic shapes. *Cricket Wind* uses graphics to provide gestural instructions, and those gestures are tracked by a sensor and used as control information for the computer sound, whereas in *Multiplexor, the composer uses* graphic notation for descriptive purposes to express sounds which are not expressible in standard notation. By using graphics as gestural notation in combination with traditionally notated piano part, the score of *Cricket Wind* promotes interplay between the pianist and computer, demonstrating conversational aspects. The two staff systems in my score will be discussed further in chapter 3.

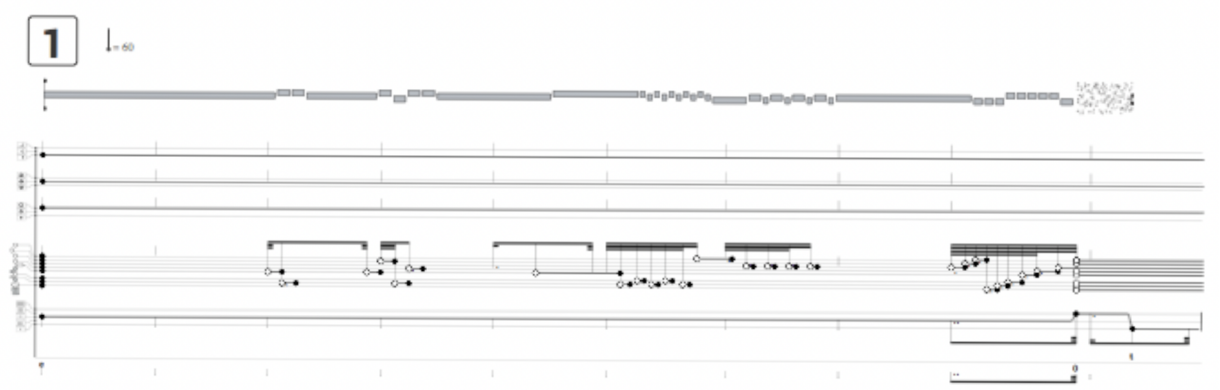


Figure 2.4 Score excerpt, *Multiplexor* by V. Adan

Works involving interactive programing

In this section, I will discuss some of the pieces written for Disklavier or involving interactive programming that are relevant for my Disklavier pieces.

Duet for One Pianist (1989) by Jean-Claud Risset

Duet for One Pianist (1989) is one of the first implementations of live interaction between a pianist and a computer that was realized in a musical context explored in short pieces (Risset 1990, 15). The title of this piece itself represents a dialogue between the pianist and a computer. By studying the work of Risset, I will analyze how interactivity is implemented in the music and what techniques I have learned and used in my own works.

Each one of these pieces, called sketches, explores and demonstrates a specific and rather simple mode of live interaction between a pianist and a computer; what the pianist plays and how

the computer responds (Rissert and Van 1996, 69). In this piece, all interactions are decided in advance, and are triggered by the performer on the piano keyboard, the Yamaha Disklavier. The interaction between the piano and a computer is implemented by the MIDI protocol under which, when a key is depressed on a MIDI instrument (Disklavier in this case), a note-on message is sent out through its MIDI output port as a message packet. It includes information about what key was played and how loud it was played. All the keys have designated numbers from 0 to 127. Similarly, the loudness is designated by the MIDI velocity between 1 and 127, 1 being the softest and 127 being the loudest. Therefore, a pair of messages is always associated with each musical note: "note-on" message and "note-off" message. When a key is released, a note-on message with velocity 0 is sent out as a "note-off" message. On the Yamaha Disklavier, in addition to keys on the piano, specific MIDI control messages can be sent from the sustain and soft pedals by depressing or raising them. The behaviors of the interaction in this piece were programmed in Max—the graphical patching language written by Miller Puckette at MIT and IRCAM, in collaboration with Scott Van Duyne at the MIT media lab (Risset 1989).

In developing seamless interactive music, Risset and Duyne were seeking ways to progress the score to change the process of operating at any given time, without the disruptive pressing of buttons on a second controller keyboard, or the clicking of the mouse. To do so, they first implemented Max's built-in module "follow" which eventually was replaced by a variety of modules—triggering pitch patterns (pitch module), dynamics (velocity detecting module) and tempo (tempo indicator)—created by Duyne for practical reasons[3]. In the case of pitch

[3] They decided to rely on Duyne's modules instead of the built-in "follow" module for two reasons: first, the compositional development was experimental hence the score did not exist but for the "follow" module to work, it needed a complete pre-defined score index number. Second, the "follow" module only takes the actual order of the notes played by the pianist into account but not dynamics and tempo (Risset and Duyne, 67).

triggering, because one-note triggers were too limiting musically, a variety of pitch modules were developed to trigger incoming note sequences of varying length. For example, when a new note arrives, a set of notes are banged out to be checked for a match. A match would then be used to trigger the next event in the piece. Similarly, the loudness can be used to trigger various events. For instance, the velocity detecting modules listen to the loudness of the last several notes played and calculate a running average, which then could be used to control parameters of a process or to trigger another event. Tempo indicators were created using Max's "timer" object to compute the times between consecutive note onsets and calculate a running average taken over a given number of the last few notes. In addition to these modules, "listener" and "playback" modules were developed to expand the compositional approach. The listener module can listen to incoming MIDI notes and store useful information about them in tables such as the key number, velocity, the duration and the offset time (the time between onsets of adjacent notes). The stored information will then be used to playback through the playback module with altered parameters. Manipulations of the key numbers can create transpositions or a variety of other pitch alterations on playback. Similarly, the loudness and tempo data could be scaled by manipulating the parameters (Risset and Duyne, 68).

Taking the advantage of live interaction between the piano and a computer, this piece *Duet for One Pianist* was composed to demonstrate various processes of live interaction between the performer and computer. In these eight sketches, a specific mode of interaction is explored and demonstrated as below (Risset 1990, 18-19):

> Doubles. The pianist plays alone, then on the repeat the computer adds ornaments. These are specified in the patch or prerecorded: they are called when the pianist plays certain notes; their tempo can be influenced by the tempo of the pianist.

Mirrors. Each key played by the pianist is echoed by a keystroke, symmetrical with respect to a certain pitch - a process used in Webern's second Variation opus 27, quoted at the beginning (and also at the end with time reversal). The symmetry center and the response delays are changed during the piece to vary effects.

Extensions: To the arpeggios played by the pianists, the computer adds additional notes transposed in pitch.

Fractals. To each note played, the computer adds five notes spaced approximately - but not exactly - one octave apart. Thus the pitch patterns played by the pianist are distorted in strange ways: an octave jump is heard as a semitone descent.

Stretch. Pitches are added, as in *Extensions,* but the intervals are not merely transposed: they are stretched by a faster ranging between 1.3 and 2.7. This stretches the chords as well as the melodies played by the pianist.

Resonances. At the beginning and the end, the computer plays long sustained chords. In the middle section, the pianist plays mute chords: the strings are set in resonance by the sequences played by the computer.

Up Down. Quasi-octave arpeggios are triggered by the pianist, whose few notes can thus generate many notes. The tempo of the arpeggios is set first by the tempo of certain patterns played by the pianist; later by the pitch he plays; then by the loudness.

Metronomes. This begins by a short canon: the computer echoes the pianist on transposed pitches and at different tempos. It later played simultaneously different sequences at different tempos. Then it repeats the same pitches, but again at different metronomic tempos, either preset or set by the pianist.

In my own Disklavier piece, *Phasing*, the central technique used throughout the piece is taken from *Mirrors* from the eight sketches. By designing to playback on the other side of the symmetry center after a keystroke, it mimics an interaction between two people. With delays, it can also vary the response time. Short or no delays (with built-in 500 milliseconds of delay time) can create overlapping effects while longer delays can create smooth transitions in turn-taking.

Gradual change of delay time could also illustrate a speaker or hearer changing their status of engagement by withdrawing gradually.

Natural Machines (2019) by Dan Tepfer inversion

Natural Machines (2019) is a project of eleven pieces for Disklavier by the pianist, composer and programmer Dan Tepfer who explores the intersection between natural and mechanical processes in music (Tepfer 2019). In each piece, one specific musical gesture is programmed by SuperCollider, accompanied by a visual representation programmed in Processing. As Tepfer improvises at the piano, the programs interact with his playing in real-time, both musically and visually. Similar to Risset's Disklavier piece, each episode of *Natural Machines* demonstrates one specific mode of interaction. What is distinct from Risset's approach is that taking the algorithm, Tepfer improvises using that as a real-time compositional device and develops his performance solely through interaction with the computer. In other words, the whole performance is built up by the interplay and conversation between Tepfer and a software program he created. I have analyzed several of his works to survey the types of musical gestures he has used, and to examine if they could be considered a representation of conversational scenes.

Ep 1: *All The Things You Are / Canon at the octave*

In this piece, incoming pitches are transposed up an octave and played back with a slight delay, which creates an exact canon of any melodic lines played by the pianist. This could

represent a situation in conversation where someone rephrases what the speaker said. However, because the phrases will be played back immediately, the canon overlaps with the melodic materials continually created in real-time, which is in a way, talking over someone in conversation. The exact canon at an octave does not conflict with the new phrases because of the perfect octave interval in higher register which is immediately played back before the new materials move on too far ahead. In conversation, when someone rephrases what was projected by the speaker, it usually happens after the utterance of the speaker is completed, not before it ends, otherwise it will become an interruption. Thus this idea of canon at the octave may not exactly fit into a realistic situation in conversation. On the other hand, the idea of playing back at an interval with delay can represent a number of situations with slight modifications. For instance, playing back without overlap could be understood as someone immediately rephrasing what the speaker said. Alternatively, the software could be programmed to randomly alter the pitches upon playback, which could represent someone rephrasing the speaker's statement in his or her own words. To purposely create overlaps, the pitches to play back should sound disturbing to some degree, hence the pitches can be randomized and can be played back before the end of the phrase.

Ep. 4: *Looper Computer*

In this piece, the computer plays back the musical phrases played by the pianist, and loops them. New phrases are layered on top of previous materials creating a thicker texture. If we consider each musical phrase an utterance from a participant in conversation, it can create a situation when the participants begin to talk over the speaker.

Ep. 5: *Inversion*

In this piece, any incoming pitch is transposed by an axis and played back with a slight delay, creating a mirror image of a pitch or melodic materials. For instance, pitches in the higher register far from the axis will be played back at same intervals from the axis on the opposite side of the keyboard in the lower register. Because the played back pitches are different as they are inverted by the axis, but are exactly the same in the rhythm, this can represent a situation in conversation in which someone responds back by rephrasing immediately after the utterance of the speaker while keeping the same momentum. This technique is exactly the same as Risset's *Mirror,* and is the primary technique I used in *Phasing.*

Ep. 6: *Intervals / Industrial*

In this piece, a single note triggers a series of pitches, 10 notes, at set intervals in a specific tempo. If another note is played while the previous set of 10 notes are being played, it creates layers of 10-note rows. The pattern of the 10 notes is changed during the performance. This idea of using one pitch to trigger a series of consecutive notes can be an interesting representation of a scene in conversation where a very short utterance such as uh, um, but, yep, etc, triggers someone to make an utterance almost like a bullet of consecutive notes. I have used this idea in my other Disklavier piece *Streams of Talk* which I will discuss in chapter 3.

Use of conversation in musical composition by other composer programmer

A Conversation-based Framework for Musical Improvisation (1994) by Walker

Turn-taking in conversation has been previously used as a framework of computer algorithms for musical improvisation by William Franklin Walker in his Ph.D dissertation *A Conversation-based Framework for Musical Improvisation* (Walker 1994). The purpose of the framework was to demonstrate how computers can participate in musical improvisation, and how the study of conversation can improve that participation. In this dissertation, two musical improvisational domains are discussed: First, a free-form three part canon defined by Sal Martirano's Sound and Logic system; and second, the tradition of small jazz ensemble performance made popular in the 1940's and 50's, with application of ImprovisationBuilder—an object-oriented framework written in Smalltalk-80 that supports models for musical improvisation. In this research, Walker applies the turn allocation system CA established by Sacks et. al to jazz improvisation by regarding solos as turns and musical phrases as turn construction units (Walker 1994, 24).

Throughout this dissertation, Walker investigates various applications of algorithms in a conversational-based framework. However, the only area of CA discussed and applied in his research is the interactive aspect of turn-taking and no other situations of conversation seem to be involved.

Chapter 3: I Hear You. Do You? — Original musical works based on Conversation Analysis

Application of conversation in musical compositions

This chapter discusses how the above theories are used and imagined in my compositions for ensemble and in solo works. I also explain how I created the conversation-driven scores which promote musicians to engage in the process of creating music through vigorous hearing and speaking.

This diagram in Figure 3.1 demonstrates how I mapped the *Nine Types of Human Interruption* by Kane according to the intentions, actions and their responses of the participants in conversation which has guided me throughout the composing process in determining and identifying the types of actions or responses needed in the musical context. For instance, to make an event of turn-taking in a piece, the first thing to determine is whether to allow overlaps or not during an interaction between the performers and how many performers are involved. If no (or little) overlaps are to be allowed, then I create a situation of smooth transition in the musical context by implementing moments for them to speak (to play) as well as transition points (during rest or at the end of a phrase). The next thing to consider is the ordering of turn-taking allocation according to the rule set 1a, b or c. If, on the other hand, overlaps are involved, the intent of the performers will determine the types of overlap to follow. If the performer enters with a competitive attitude intending to disrupt the current ‘speaker’ or to take over the turn, it will yield overlaps under “competitive” in the diagram. Specific types of competitive overlap will

then be determined by the action of the current speaker who was disrupted. If the current speaker stops talking (or was forced to stop), options are either competitive overlap or competitive pause overlap. If the current speaker continues despite being interrupted, then it can be either pause butting-in or speaking butting-in depending on whether the both the original speaker and the person who interrupted talk or only the original speaker talks. If the overlap is created with an intent to support or agree with the current speaker, then options are collaborative overlap or collaborative pause overlap if the current speaker ends up passing the turn to the person who interrupted or has to make an adjustment. If the current speaker is able to continue, then the interruption will be a form of backchanneling. When an overlap occurs unintentionally, i.e. both people start talking at the same time after a silence, it is neither competitive nor collaborative, therefore, will be considered an accidental overlap. By the time the type of interruption is determined, the surrounding situations and context will have been designated already. Then, I as the composer can translate the scenario into a musical context by using it to guide the structure of the piece. In particular, I would consider what kind of musical gestures can imply competitive, collaborative or neutral intent by asking myself these questions: If overlapped, what kind of harmony would sound collaborative and supportive in contrast to competitive? If both parties decide to continue speaking after an overlap, how will I want their rhythms and harmonies to interlock or not?

With regard to displays engagement and disengagement, the questions in the box of Figure 3.2 are what needs to be concerned with to identify the phases of engagement or disengagement. These behaviors, actions, intentions and body orientations, whether explicit or implicit, can elicit varying displays of engagement, requiring different musical interpretations.

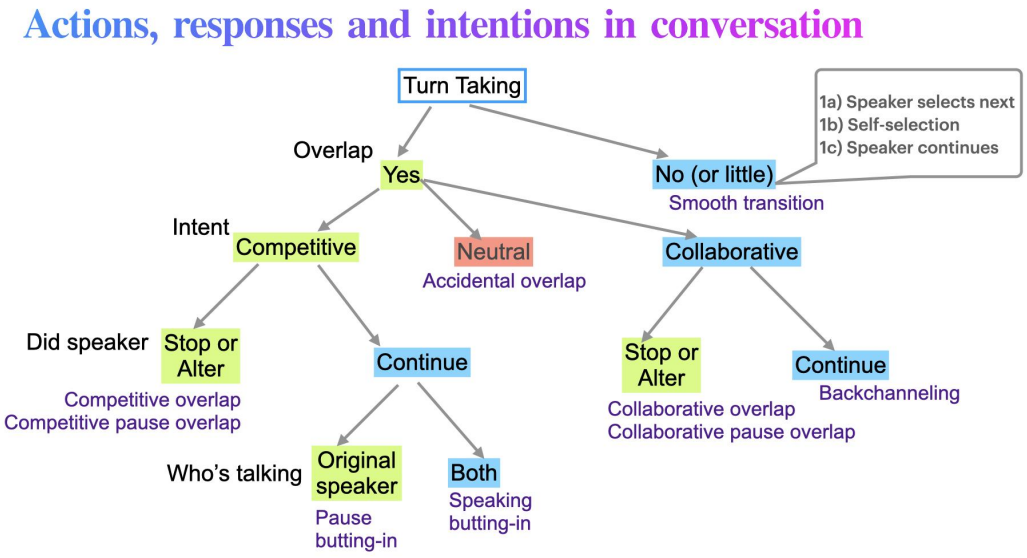


Figure 3.1 Map of the *Nine Types of Human Interruption*

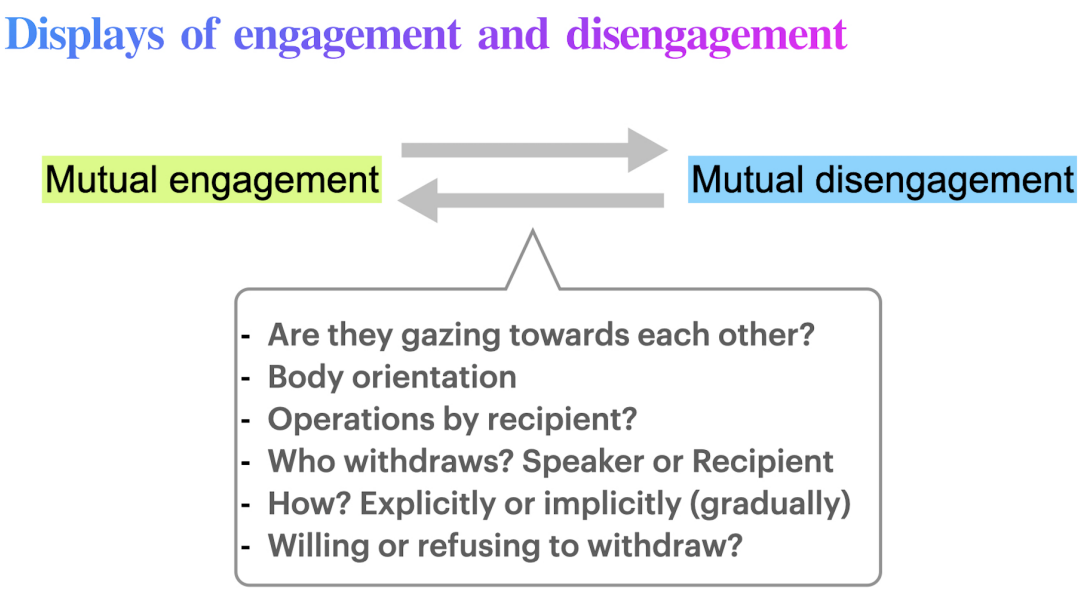


Figure 3.2 Displays of engagement and disengagement

Group Talk No.1 - 3 (2021)

Members of ensemble

Group Talk (2021) is a suite of three pieces which model possible situations in a conversation by a group of individuals who come from different cultural and musical backgrounds. Musicians in this ensemble are Blake Harrison-Lane on violin, Bella Pepke on cello, JoVia Armstrong on percussion, and myself on piano, who are each gifted with different strengths, skill sets, and aesthetics in music. The violinist Blake and the cellist Bella are primarily trained in traditional Western classical music but are both familiar and proficient in improvising in contemporary music settings with or without extended techniques for strings. Bella is more performance-oriented, and has also experience in playing jazz, whereas Blake considers himself to be more a composer than a performer. Nevertheless, both are highly skilled and proficient, capable of reading Western notation as well as other contemporary or free-form scores for improvisation. JoVia is an experienced percussionist whose strength is the ability to sense the essence of the music with only a glimpse at the scene, and the versatility to adapt immediately. As participants in conversation are constantly monitoring and adjusting their actions, JoVia's ability to observe and comprehend the situation instantly and adapt to any changes will be highly valued in this ensemble. She is trained mostly in improvised music such as jazz and other contemporary music, and is able to read notated music. What sets her apart from other performers is that she has experience working in hybrid groups in which musicians come from diverse backgrounds including jazz, classical, and traditional Japanese music. I believe that is why she has the ability to create and maintain a cooperative atmosphere when

working with musicians from different practices. I consider myself to be an in-between player who was first trained in Western classical music and later in jazz and other forms of contemporary music. Though I feel comfortable in either setting, I feel the most comfortable when both the music and the people are a fusion of multiple cultures. During the performance with this group, my role is as a bandleader as well as a pianist.

Workshops with musicians

Prior to having rehearsals, I had the opportunity to work with the string players from November 2019 in the form of improvisation workshops. In doing so, I had two primary purposes in holding these workshops in mind. First was to experiment with compositional ideas with the members of the ensemble, Bella and Blake, which not only helped me hear the sonic outcomes of my ideas but also helped me understand the limitations, constraints as well as capabilities and possibilities on the instruments. Second was to practice different improvisational ideas. Below are examples of the practice games we workshopped (Figure 3.3). Whether or not these workshops contributed to improving their skills, these exercises seem to have overcome hesitancy and cultivated willingness to contribute or express by creating their own melodic lines.

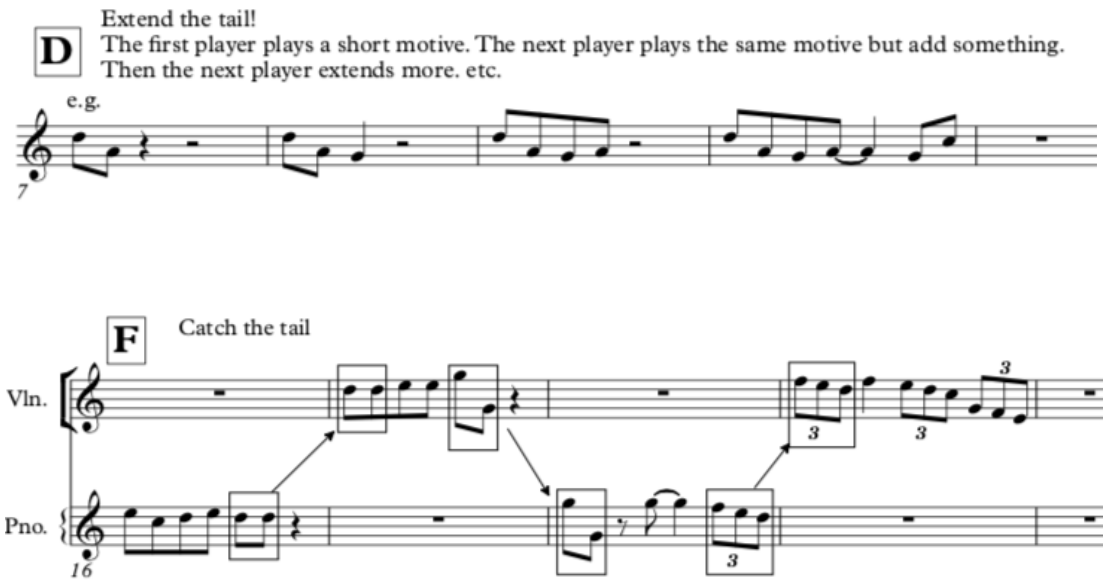


Figure 3.3 Excerpts of exercises from workshop

Conversation Analysis in *Group Talk no. 1-3*

This suite of music consists of fully notated precomposed sections and sections where only a framework or instructions are provided for the performers to improvise. While these precomposed and improvised sections interweave throughout the suite, a number of possible conversational scenes are exemplified in the music. In the improvised sections, performers are granted the choice to select materials to play from options, or to improvise within the framework provided. Each scene is called, develops, or advances to the next section by sonic cues initiated from a member of the group.

Below are the types of scenes in conversation manifested in the music:

- Turn-taking.

- Different states of engagement and disengagement: observed from 1) operations by recipient during talk, 2) visible withdrawal from talk by speaker, 3) recipient withdrawal, 4) activity-occupied withdrawal, 5) refusing to withdraw, 6) maintaining availability, 7) matching displays, and 8) talk with disengagement.
- Human interruption defined by Kane: collaborative overlap, competitive overlap, competitive pause overlap, speaking butting-in, and backchanneling.

Group Talk No.1

Turn-taking and Accidental overlap

In section A, the interaction between the violin and the piano exemplifies turn-taking and accidental overlap[4], as they make utterances in the form of short bursts of sounds behind the cello playing continuously sliding dissonant pitches as shown in Figure 3.4. This turn-taking expects smooth transitions between the two as well as accidental overlaps at times because they are interacting in a free-form manner in terms of timing. One can choose to wait for the other party to make the next utterance, or one can continue to talk with little or no gap in between for the other to take the turn. As the intensity of the background sound by the cello increases, the interaction between the two also intensifies, causing more chances of overlapping. However, because the utterances in this section are short, the moments of overlap are short and brief without being completely taken over.

[4] Accidental overlap - After a prolonged silence both people might start talking at the same time, but eventually one of them will give up and one carries on the conversation (Kane).

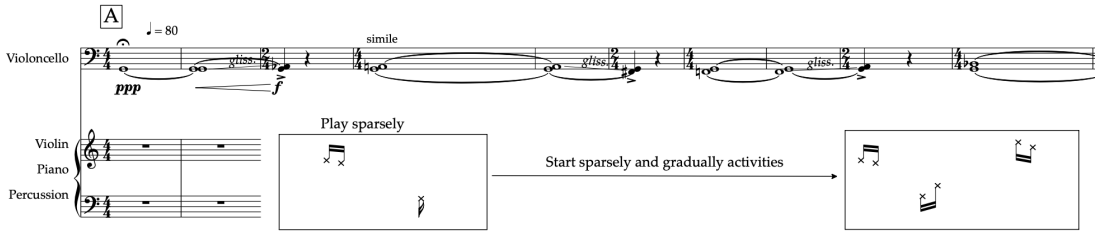


Figure 3.4 Opening of *Group Talk No.1*

Mutual disengagement and Speaking butting-in

Mutual disengagement is recreated from measure 38 where two parties are co-present, but not engaged in the collaborative activity yet both are observing each other. From measure 38, the violin stops exchanging utterances with the piano and begins to play a series of ascending whole tone scales in sixths (Figure 3.5). During this time, the piano continues to talk freely with increased intensity as if she does not hear or intentionally chooses not to react to what the violin is playing. In this scene, both parties are being the speakers and talking with disengagement. Neither of them seems to place an expectation to be heard by the other party, hence both speakers show no signs of indication that allows the other party to respond, and create continuous utterances. On the other hand, this is also an example of speaking butting-in[5] as a type of human interruption as seen in the violin's behavior from measure 38 continuously playing without giving a space for the piano to respond or paying attention but rather being competitive towards the piano.

[5] Speaking butting-in is one of the nine types of human interruption defined by Kane. One person is talking and another begins with the intent of taking over the conversation, however, the original person continues to talk regardless.

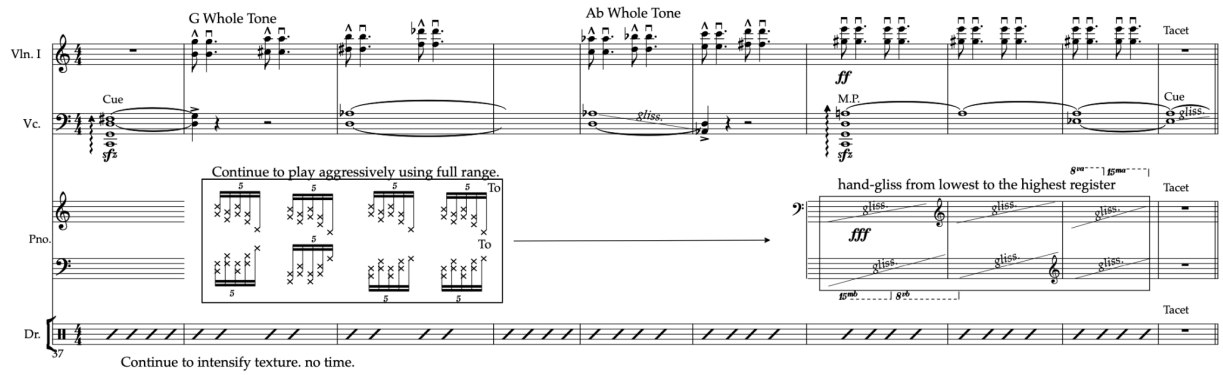


Figure 3.5 Measure 37 of *Group Talk No.1*

Accidental overlap

The last section exemplifies a group conversation that gradually turns into accidental overlap. All the players will attempt to play on the downbeat of one. However, whenever one tries to speak, somebody else cuts in to interrupt which prevents the speaker from continuing (Figure 3.6).

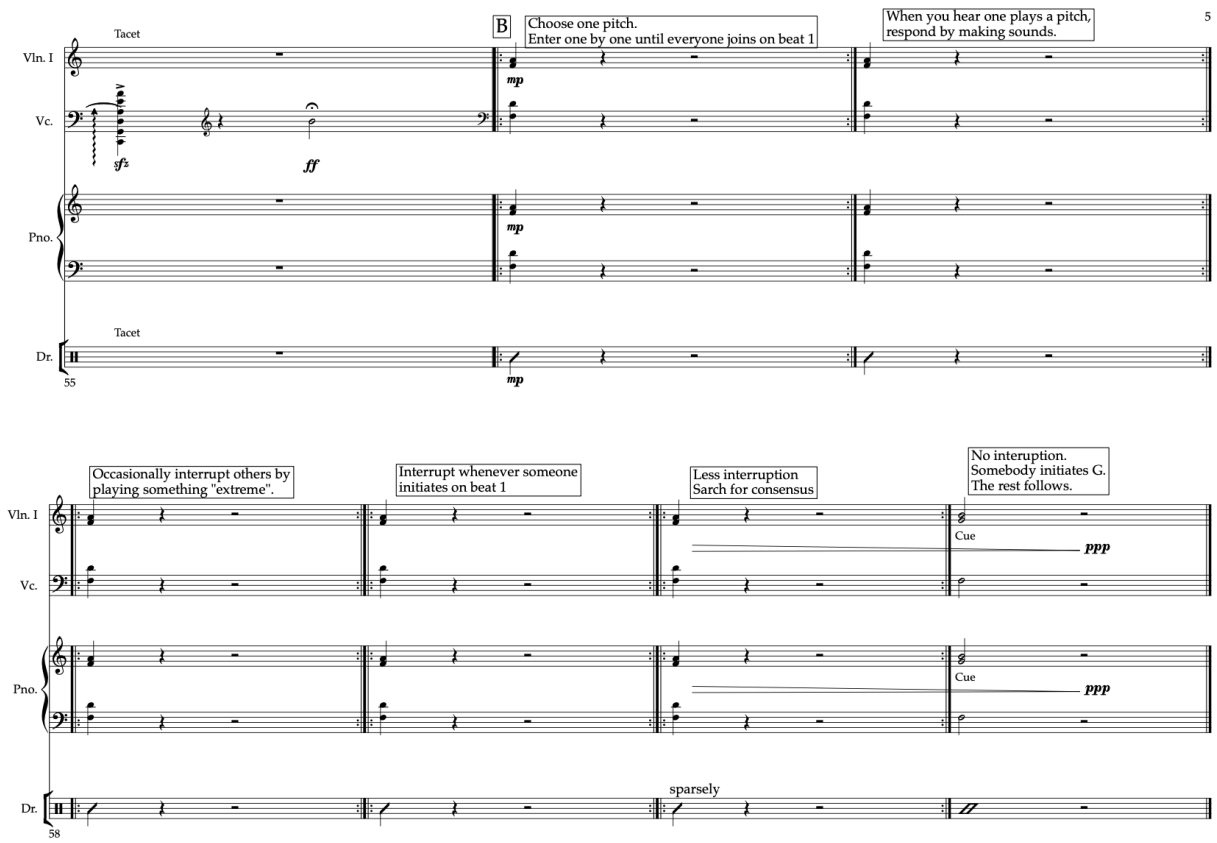


Figure 3.6 Section B of *Group Talk No. 1*

Group Talk No.2

Turn-taking and Smooth transitions

A gesture of simple turn-taking between two parties is depicted in the opening section of Group Talk No. 2. In section A, violin and cello interact and develop a dialogue by exchanging pitches on assigned beats with a specific pitch; any pitch A on beat one for cello and any pitch D on beat three for violin. Within this framework, they have the freedom to orchestrate or embellish their pitches in their own ways by changing the registers, doubling in octaves or with

various articulations. While each utterance is only one beat long, it exemplifies a smooth transition[6] as this turn-taking is performed without interruption and overlap (Figure 3.7).

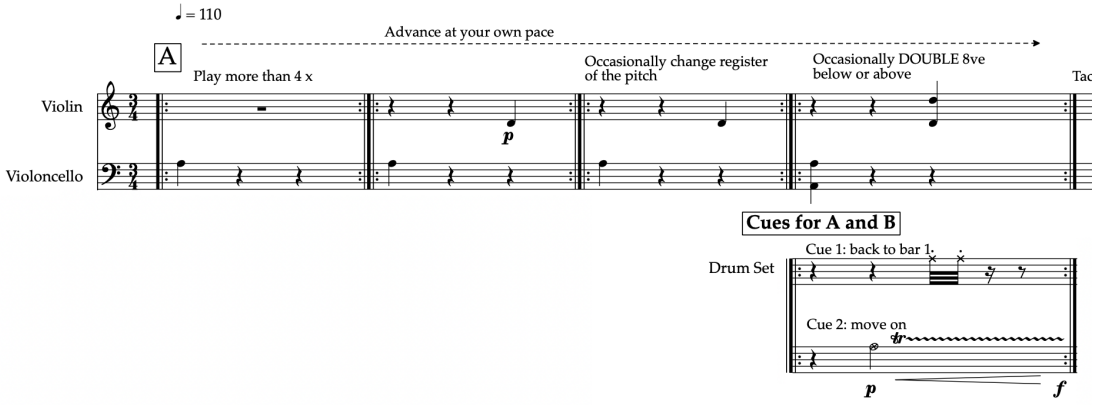


Figure 3.7 Section A of *Group Talk No. 2*

In this section, I have inserted an external interruption. While this is not a type of human interruption that arises from one of the participants of the conversation, at any point during conversation, any kind of interruption or distraction can occur in actual situations whether it be a phone call or siren from outside. When the cue is called, the cello and violin have to return to the very beginning and start the whole process all over again. The intent of injecting such an interruption is to create a situation in which the participants are forced to pause and make them start all over again because the interruption either ruined the flow of the conversation or made them forget where they were. Cue 2, on the other hand, functions as a cue to advance the music to the next section.

[6] Smooth transition - Both speakers are able to finish their turn, without interruption (Kane).

Operations by recipient during talk and Backchanneling

From section C, the staffs previously assigned to the violin and cello are now labeled as player 1 and 2 (Figure 3.8). The intent is to allow them to be either by taking an initiative. Whoever initiates the consecutive eighth notes in measure 14 will be the player 2 for the first turn, being the hearer in the next eight bars starting measure 15, and the other player will automatically be the player 1 as the speaker from measure 15. The gesture of player 2 displays a type of operation by the recipient such as nodding or a short utterance such as “yeah”, “uh-huh”, or “really” known as backchanneling or active listening with the intent of showing attentiveness to the speaker.

Figure 3.8 Section C of *Group Talk No. 2*

Turn-taking 1a), Smooth transition, and Backchanneling

Transitioning from section C, section D demonstrates turn-taking rule 1a) starting as a smooth transition coupled with backchanneling. During the last two bars at the end of 8-bar cyclic cycle in section C, one of them will create a bar-long rhythm to be repeated throughout the next cycle while the other party will respond with one utterance on upbeat of three as a form of backchanneling (Figure 3.9). The current hearer will take the next turn by establishing a new rhythm at the end of the section, and will be the speaker for the next cycle. This allows the two to speak without interruption on their turn while having the hearer being present and attentive as they make smooth transitions. Although the current speaker does not explicitly say or give a cue to pass on the next speakership to the current hearer, the basic premise here is that both participants take turns as rule 1a).

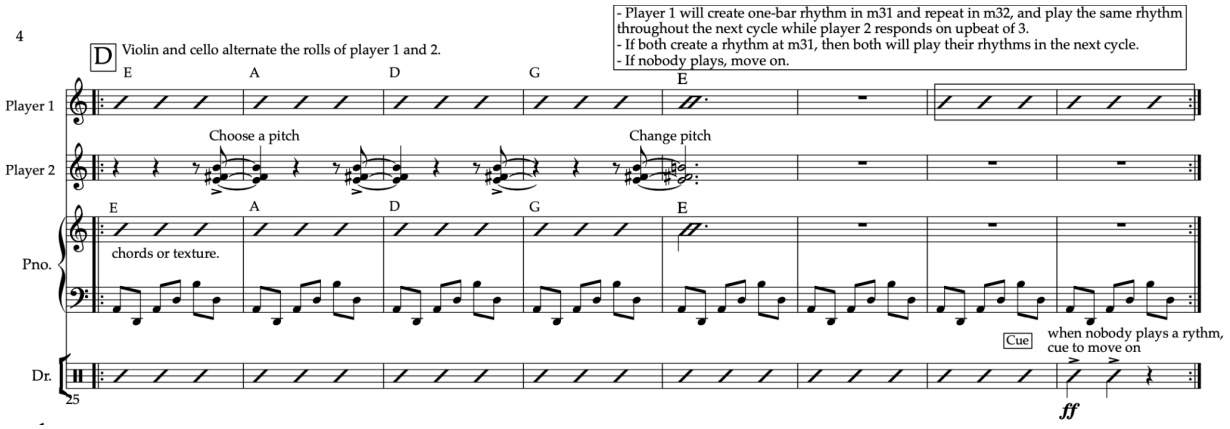


Figure 3.9 Section D of *Group Talk No. 2*

Turn-taking Rule 1b) & c), Accidental overlap and Collaborative pause overlap

Beside smooth transition through turn-taking rule 1a), section D also exemplifies other possibilities such as rule 1b)[7], c)[8], collaborative pause overlap[9] and accidental overlap. While this section presumes the two to take turns every eight bars, it also assumes odd situations in which one may choose to be the speaker again or opt out for being the next speaker after being a hearer and remain silent. If the current speaker creates the next rhythm again in the last two bars of the cycle instead of passing the turn to the hearer, that would cause an accidental overlap, which also implies a kind of self-selection where the current speaker chooses herself. If such an overlap happens in an actual conversation, usually the parties will pause to figure out who should take the turn. It is rare in actual conversation that this kind of overlap is left uncorrected unless two parties are arguing, and none of them[10] is willing to negotiate. However, if this overlap happens during the performance, the players can carry on with the rhythms they established in the next eight bars. As a type of interruption, this would imply collaborative pause overlap since the next 8-bar cycle will sound neutral (neither being competitive or disagreeing) and collaborative as they create contrapuntal lines being cohesive to the given harmony. In the turn-taking rule set, if nobody selects to be the next speaker, then rule 1c) applies, thus the current speaker has the

---

[7] Rule 1b) If the 'current speaker selects next' technique is not involved, then self-selection for next speakership may, but need not, be instituted; first starter acquires rights to a turn, and transfer occurs at that place.

[8] Rule 1c) If the 'current speaker selects next' technique is not involved, then the current speaker may, but need not continue unless another self-selects (Sacks, Schegloff and Jefferson, 704).

[9] Collaborative pause overlap - A person may pause during a conversation and wants to continue, but before they can a second person starts talking. Here, the original speaker lets the other person continue, making the interaction sound collaborative or neutral.

[10] Rule 1 c) If the 'current speaker selects next' technique is not involved, then the current speaker may, but need not continue unless another self-selects (Sacks, Schegloff and Jefferson, 704).

choice to continue speaking or not. Instead, in this music, this accidentally produced silence is used as a cue to advance to the next section.

Entering state of engagement and disengagement

Section E exemplifies how a group of participants enters a state of engagement and gradually shifts to a state of disengagement, which is represented by participants assembling one by one to join the conversation, and leaving gradually as they slowly withdraw (Figure 3.10).

The players were asked to create phrases using the given pitches, and develop by adding harmony. They will enter one by one creating their rhythms which they individually vary in their own ways, which will contribute to developing the music collaboratively as a group. These gestures imply a scene of group discussion where one initiates a talk, then a second person, until the fourth person joins and everybody is in full engagement to the ongoing conversation by actively exchanging a form of utterances whether it be a sentence or a sound to display the attentiveness.

Once the music culminates, they start to shift their attention away from the ongoing conversation taking place in the same space, as they enter the state of disengagement gradually, which is performed by removing one note at a time, becoming more sparse.

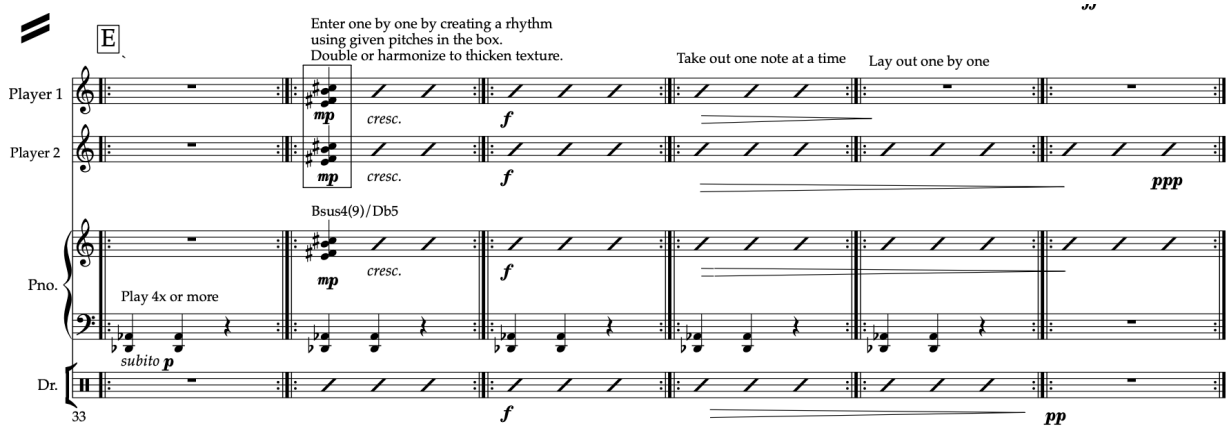


Figure 3.10 Section E of *Group Talk No. 2*

Group Talk No. 3

Collaborative overlap through Turn-taking 1b)

Group No.3 explores conversation between not only two participants but also four participants. Section A demonstrates an abstract image of collaborative overlap[11] between four participants in which all four initiate utterances at their own will and pace, beginning quietly in the mid range and gradually expanding the range towards the higher register and lower register along with dynamics and intensity (Figure 3.11).

Having no set time or order of who begins first and goes next, the resultant effects in music may be unorganized frequent overlaps. This could easily lead into a cacophony full of chaotic and competitive overlaps[12] and interruptions, were it an actual conversation. Instead, the

[11] Collaborative overlap - one person starts talking before the other finishes and because of this, the person who was speaking first, has to after what they were going to say or even stop speaking prematurely (Kane).

[12] Competitive overlap - one person starts speaking before the other finishes and the interruption is not welcome.

overlaps are converted to a more cooperative and congruent sound in music by maintaining the same harmony, D lydian mode, throughout the section. All the players will be choosing notes from D lydian mode to play trills or fast tremolos. Piano and percussion will imitate the same gestures and adapt to their instruments. The resultant sound will be emerging swells of sound arising from the four players during which they overlap with each other creating layers of swells. Allowing all the players to initiate on their own also exemplifies self-selection, rule 1b) in turn-taking except that there may be more than one player or party speaking at a time.

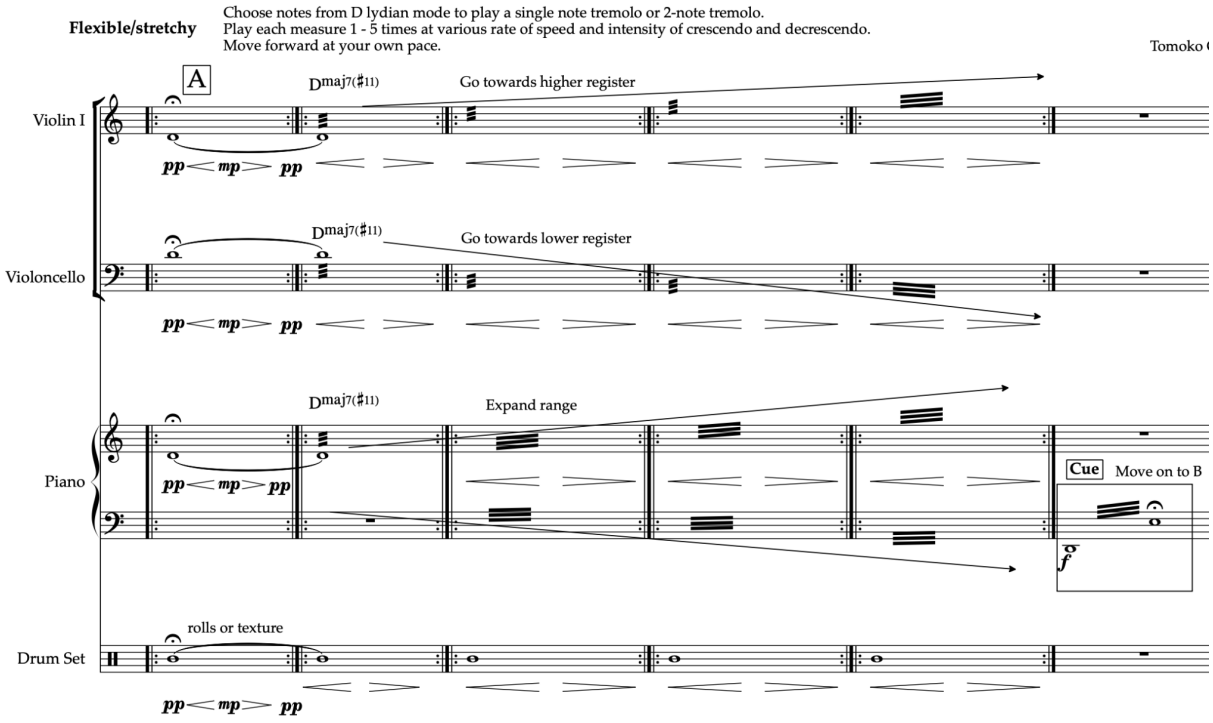


Figure 3.11 Section A of *Group Talk No. 3*

Various types of withdrawal: Visible withdrawal from talk by speaker, Recipient withdrawal, Refusing to withdraw

The scenario of section B is a cocktail party conversation in which players are given five different examples of shapes that they are able to freely choose and move around after staying with a shape for more than a certain number of repeats. In the same way as section A, the harmony of section B is set to D lydian mode (Figure 3.12). This will ensure the sound to be cohesive in terms of harmony even when players freely move between the shapes and choose different pitches, causing unintentional overlaps and harmonies at different intervals. This, in return, will create a blend of evolving sounds as a result of different combinations of the shapes, as well as different harmonies within the same shape.

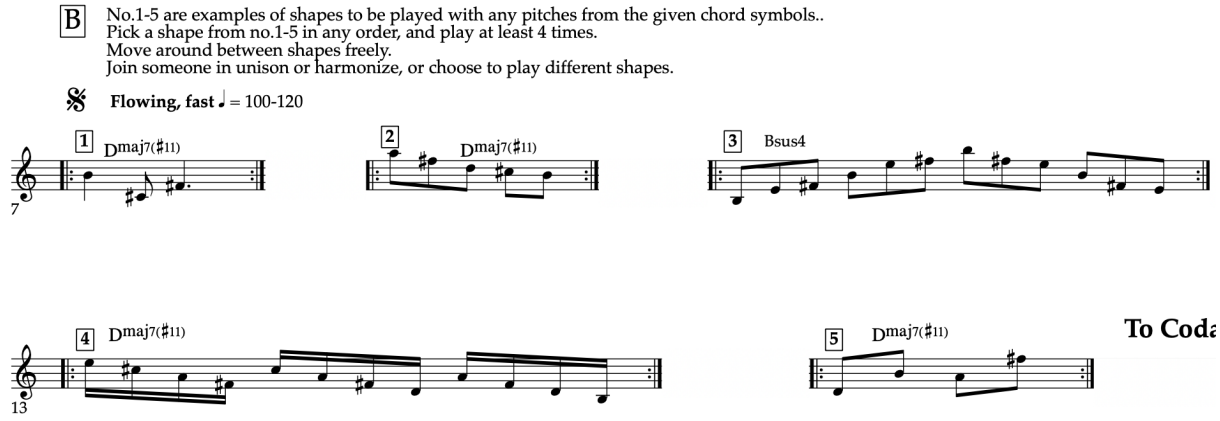


Figure 3.12 Section B of *Group Talk No. 3*

The setting of a cocktail party in this section allows players to join a conversation for a short while, and leave to join other conversations at any point. They are expected to play the same shape more than four times before moving to another shape, not only for them to settle in but more importantly to let other players notice, and give time to react or not react. If we

consider joining or staying with the same shape a form of engagement, opting out from a shape to start another shape will be considered withdrawal entering a state of disengagement while being physically present in the same space. Even during the state of disengagement, participants are monitoring each other. This is to enable them to make conscious choices as to how they want to collaborate or not. While performers experience continuously changing states of engagement and disengagement, they will be creating collaborative overlaps as well as going through stages of withdrawals such as visible withdrawal from speaker[13], recipient withdrawal[14], and activity-occupied withdrawal.[15] The only unique feature here is that all of the participants are speakers and hearers, and the roles are fluidly changing in a non-orderly manner. To withdraw, the players can gradually diminish their playing or slow down, or transform their current shape into the next shape, which enables them to withdraw from the conversation they are in while being present in the same space.

---

[13] Visible withdrawal from talk by speaker: Speaker withdraws her gaze and becomes silent, and changes position of cigarette. Though the speaker both withdraws her gaze and becomes silent, she is still actively involved in producing her talk but with less than full engagement. Recipients do not attend to the gaze withdrawal as an isolated event, but rather analyze it with reference to other activities the speaker is performing at that moment.

[14] Recipient withdrawal: Recipient begins to nod and withdraws her gaze while the speaker talks; the withdrawal is occupied by talk-relevant activities. Nods during withdrawal are not performed in the same way as those done during full-engagement; they are slower and more subtle. Though recipient's nods continue to perform actions relevant to the talk, these actions are performed in a way that is sensitive to the changes in engagement states that are occurring.

[15] Speaker continues to talk while withdrawing from her recipient and positioning herself for entry into a different activity. The boundary between full-engagement and mutual disengagement is not structured as a sharp break. Participants are given a space where they can reorganize their bodies and actions to show the change and display their understanding of the talk. The act of dis-attending each other never emerges for either party as a noticeable, recognizable activity in its own right. When each participant finishes her talk-relevant activities, she finds that she and her co-participant are no longer in orientation toward each other, that state of affairs having been systematically achieved but never made visible as an explicit act of disaffiliation (Goodwin, 106).

Turn-taking 1a), Smooth transition, Speaking butting-in and Matching displays

Another smooth transition in turn-taking is displayed in section C as a duet by the cello and percussion. The two will trade turns between measure 26 and 29 (Figure 3.13). As they trade turns, each will improvise in the measure of 3/4. The cue in measure 30 is called by whoever wants to stop the trading and overtake the turn, which is another example of speaking butting-in as the other player will continue to play the rhythm. The hit on beat three in measure 33 is meant to give an impression of attentiveness as a matching display[16].

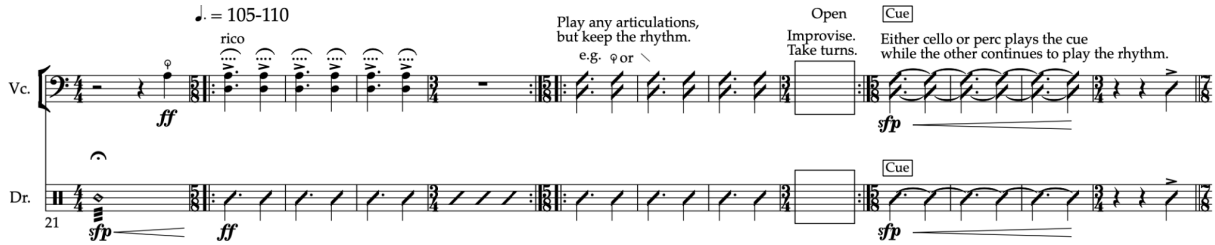


Figure 3.13 Section C, measure 21-33 of *Group Talk No. 3*

Smooth transition and Collaborative overlap

Section D is a contrasting duo by the violin and piano which exemplifies smooth transition and collaborative overlap. Basic example forms of gestures are given to the violin to play during the interaction with the piano starting at bar 58 (Figure 3.14). The piano will outline the four chords with different embellishments in each iteration, handing to the violin to fill in measure 59 (Figure 3.15). Each of them is given an opportunity to be the main speaker for a full bar; the piano takes turn in measure 58, and the violin does so in measure 59. These two

[16] Displays of eyebrow flash demonstrate attentiveness to the speaker and also that the recipient has in some way dealt with the particulars of the talk of the moment. It also allows her to perform explicit withdrawal by dropping the brows. When both the speaker and recipient perform eyebrow flash, they demonstrate a display of congruent understanding hence not providing further talk may be relevant. This explains why such actions are often found at places of transition from talk to disengagement (Goodwin, 110).

measures will repeat as they develop the conversation collectively by hearing each other and expressing themselves fully and freely as if they are having a constructive discussion. While doing so, they can continue to improvise beyond their measure, overlapping the other player. Nevertheless, because there are no competitive intentions or attitudes, the overlaps will be considered as collaborative overlap, which, musically speaking, adds a counterline on top of each other.

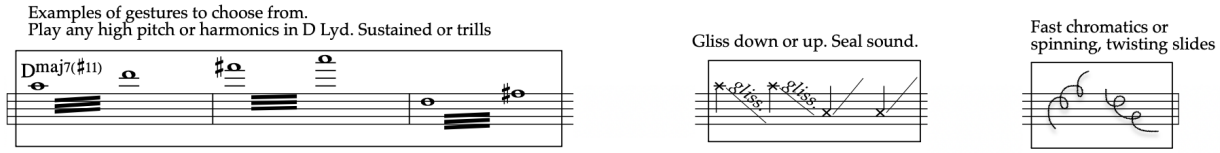


Figure 3.14 Gestures for violin in Section D of *Group Talk No. 3*

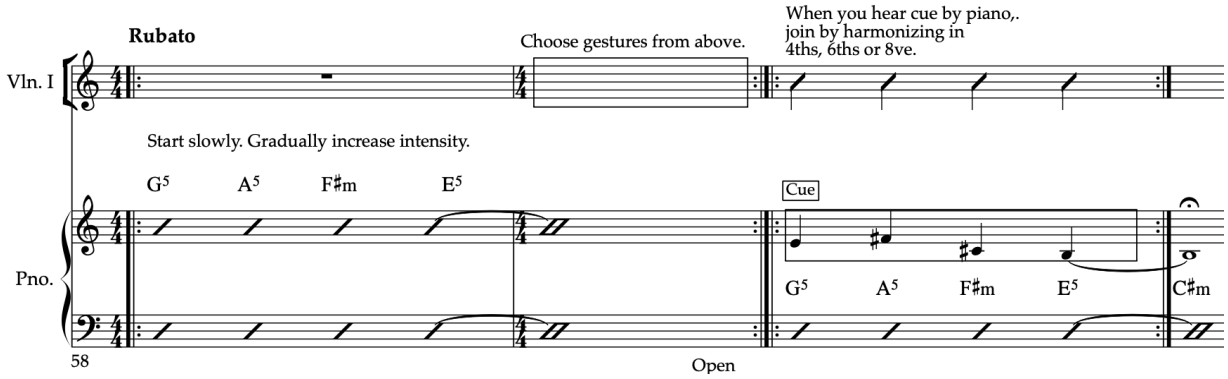


Figure 3.15 Section D, measure 58-61 of *Group Talk No. 3*

Cricket Wind

Cricket Wind is an interactive solo piano piece with the motion sensor, MUGIC™, programmed in Max for Live. The music attempts to recreate the orchestral sound-scape I experienced in the park while in Japan in 2020. In the music, I seek connection and conversation

with nature through the sound of the piano and the sample of crickets and the gestures from performance on the piano. The wind sound, which I call “cricket wind” is created by the sample of crickets which was recorded in a room of the Cricket Temple in Kyoto, Japan where approximately six thousand crickets are kept in the cage by all the walls. It is made into white noise-like wind sound by applying multiple filters and echoes in Ableton Live. The sensor in the glove worn by the right hand will control the volume of the wind by the vertical movement of the wrist. The rest position of the hand, when the wrist is lowered while the fingers rest on the piano keyboard, marks the lowest volume. As the hand tilts down and the wrist rises up, the volume will increase. This up and down motion occurs naturally as part of piano performance: when beginning to play from the rest position, during the performance while making various gestures, or when coming to stillness. In the piece, these gestures are used to generate the wind sound on its own without the piano, as well as to create interaction between the sound from the piano and the wind.

Smooth transitions, Operations by recipient and Backchanneling

While the scenario of this piece is purely imaginative, the first two sections of the piece are designed based on a turn-taking model. For instance, the opening of the piece depicts a dialogue between the piano and the wind. The piano opens a conversation with a short statement in measure 2 as the speaker (Figure 3.16). When the phrase completes, the wrist floats up and down, inviting the wind to respond which forms a smooth transition. The response of the wind is the operations by the recipient such as nodding or short utterances to display its attentiveness. As the music progresses, the playing of the piano increases its intensity making the wrist fluctuate

more vigorously, which invites a bigger remark from the recipient, the wind, as seen in measure 11 (Figure 3.17). From measure 19, the piano will be speaking in a more excitable way, making more active movement in the wrist vertically, which makes the wind follow as if the wind is backchanneling (Figure 3.18).

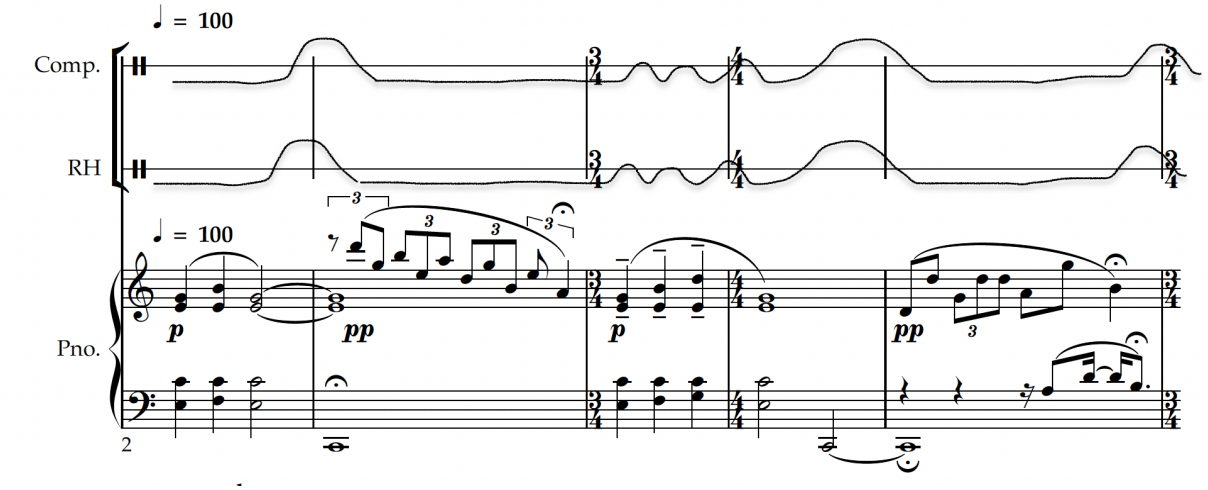


Figure 3.16 Measure 2-6 of *Cricket Wind*

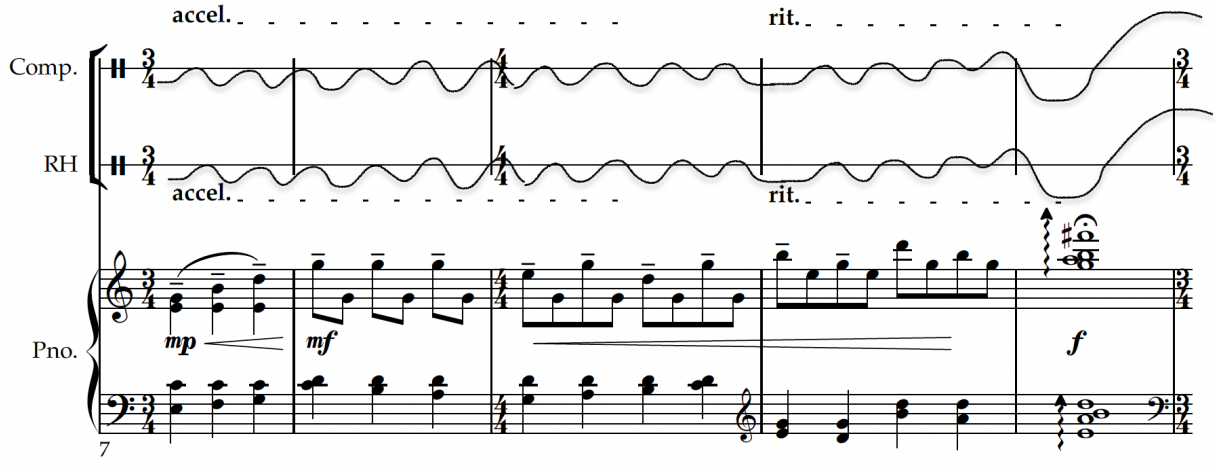


Figure 3.17 Measure 7-11 of *Cricket Wind*

Collaborative overlap

The distinction between the speaker and hearer becomes blurred as the playing and the fluctuating movement of the wrist intensify, making the sound of the piano and the wind merge together and overlap with each other (Figure 3.18). These overlaps are conducted in a collaborative way, hence exemplifying a form of collaborative overlap.

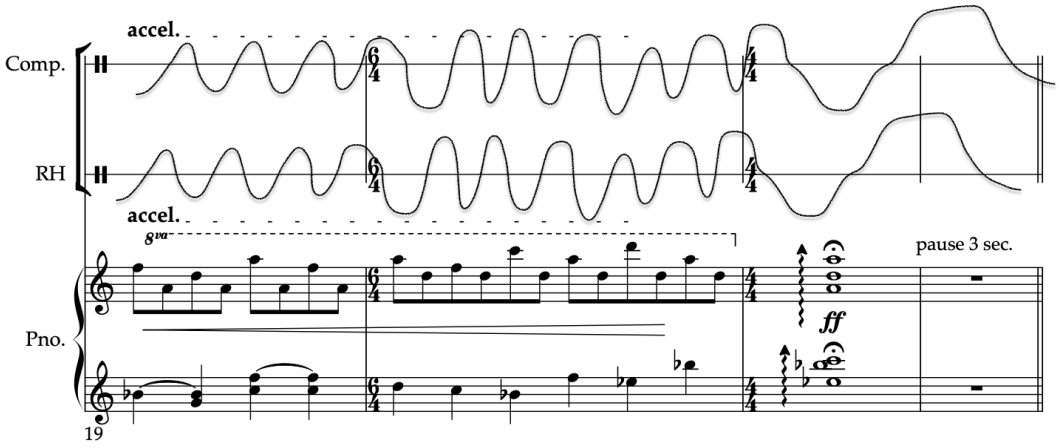


Figure 3.18 Measure 19-22 of *Cricket Wind*

Smooth transition and Operations by recipient

The interaction between the piano and computer in section B exemplifies smooth transition and operations by the recipient in conversation, which is enabled by the motion sensor reading and responding to the inactivity of the right hand (Figure 3.19). When the piano plays short bursts of sounds as utterances, the computer plays back the delayed grains of the piano sound as a response after the right hand comes to stillness. The sensor begins timing the duration from the moment the sensor comes to stillness by reading the activity level of the hand. The wet/

dry mix and the feedback levels of the grain delay, an effect from Max for Live, are scaled by the duration in stillness between 100 and 1800 milliseconds in reverse. Hence, the longer the pause, the greater the response will be in terms of feedback and the volume. Programming the computer to respond to the inactivity of the hand forces the piano to complete its utterance before passing the turn to the computer, to make smooth transition without overlap. As a conversational scene, I am imagining myself being surrounded by trees, attempting to communicate with creatures or beings in nature whether it be birds, insects, winds, trees or flowers. The pause after playing a short phrase and the irregularly delayed response depicts myself waiting to see if anybody is listening and will respond. The response of the computer is different every time depending on what the piano plays and when it stops, which offers a variety of responses as operations by the recipient.

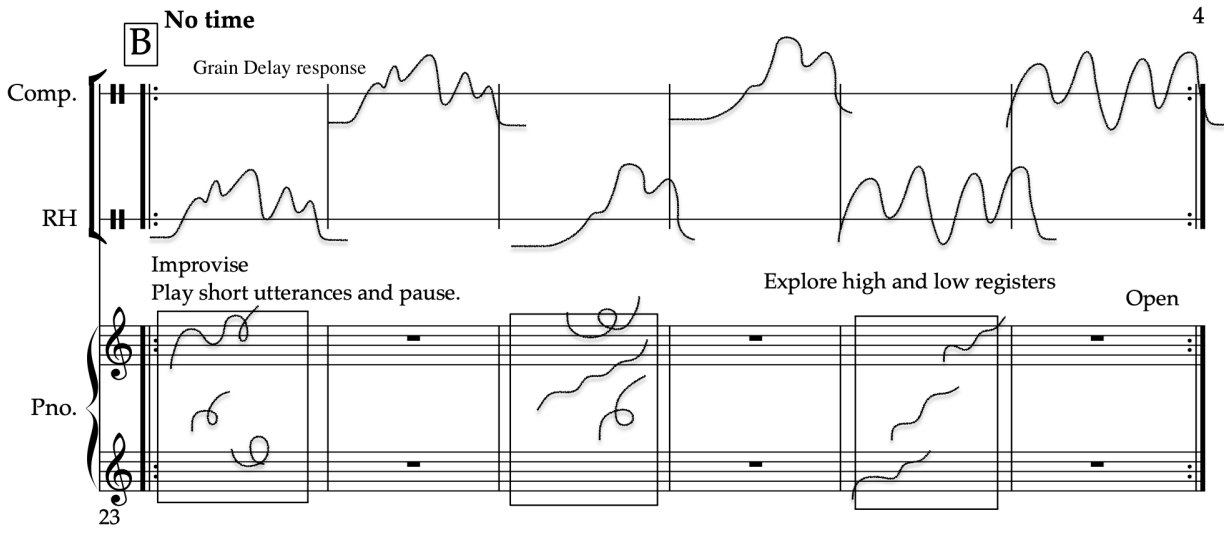


Figure 3.19 Section B, measure 23-27 of *Cricket Wind*

Phasing

The music is a solo piece for Disklavier which attempts to depict various moments of dialogue using MIDI and an interactive program by MaxMSP. This piece was initially inspired by *Piano Phase* (1967) by Steve Reich, though this piece has taken its own path. The title *Phasing* is the representation of transformation of musical and conversational scenes. As for the practical programming, the ideas are borrowed from *Inversion* from *Natural Machines* by Dan Tepfer (Sunnyside Records, 2019). The main feature of this piece which contributes to making the music more conversational is the reflection of the playing created by axes and delay times.

Conversation by two is emulated by inverting the input notes around an axis with delays. Thus, what is played on the Disklavier will be mirrored on the other side of the axis. For instance, if the input note is G4, with an axis E4 being minor third apart, the computer will transpose G4 to C#4, a minor third below E4. Since this piece is programmed to be performed with the built-in delay of 500 milliseconds on the Disklavier, the MIDI playback on the Disklavier will occur in exactly 500 milliseconds after the MIDI note message is received by default when the delay time in the program is set to 0 milliseconds. Therefore the mirrored image of the playing will always be heard with a slight delay even when the delay time is 0. In some places, the delay time is fixed to create a response or playback in a certain amount of time, allowing the Disklavier and the voice in the mirror to interact, or create an effect where the Disklavier is being chased or shadowed. In other places, delay time changes over time, displaying a gradual shift in the engagement status of the two participants, the Disklavier and the mirror. As the axes change, the mirror image of the voice is transposed to different places in the register, creating continuously evolving harmony.

Speaking butting-in

The opening of the piece is performed freely as a solo piano. As a conversational setting, this will be a speaker enjoying his turn, expressing fully uninterrupted. However, as the mirror axis is opened by the last note of bar 11, everything that the speaker says is shadowed and inverted from bar 12 with 500 milliseconds delay (Figure 3.20). Because the speaker ceaselessly talks, so will the mirrored playback, making it sound like the speaker is constantly being interrupted and overlapped, causing speaking butting-in since the speaker continues to talk regardless.

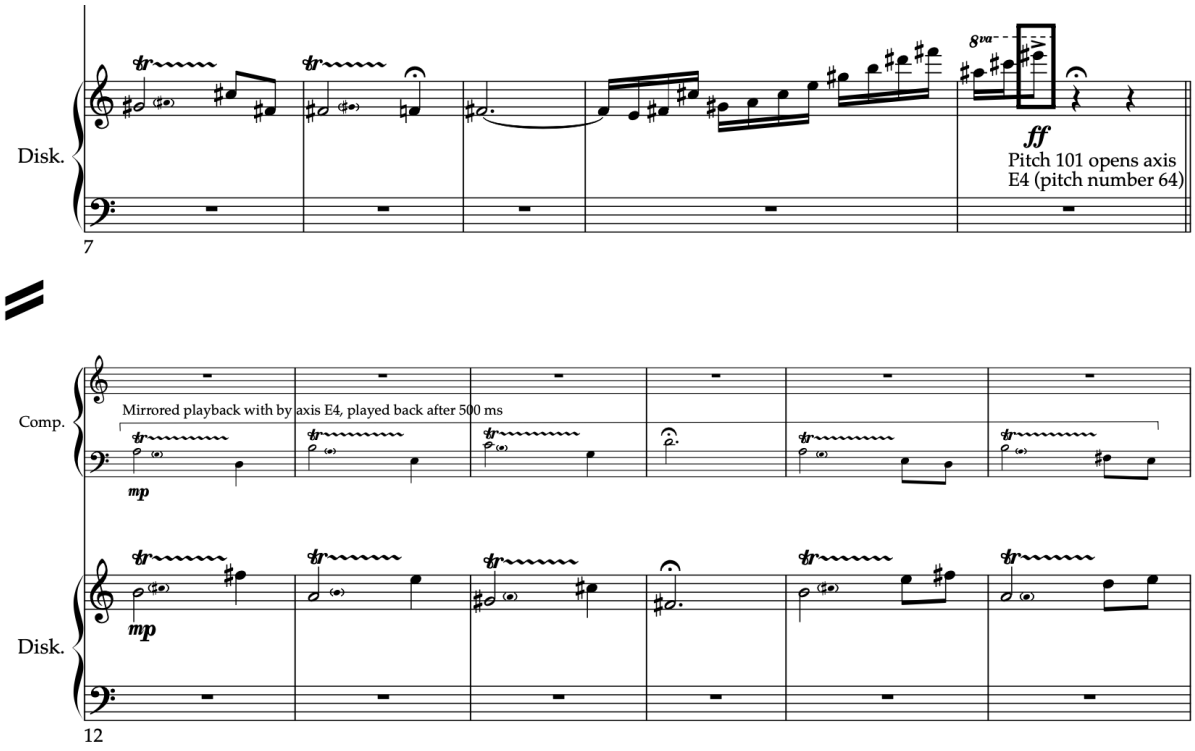


Figure 3.20 Opening of *Phasing*

Turn-taking

On the other hand, using exactly the same inversion axis with no added delay time, a different approach is taken in section B to create more interactive turn-taking between the speaker (Disklavier) and the mirror (inverted playback), than the previous section where it was more one-way type of communication. The gesture of the Disklavier here utilizes the built-in 500 milliseconds delay time to give space for the mirror to respond back (Figure 3.21). At 120 bpm, two eighth notes will make the mirror respond on the next beat which will create a call and respond type of interaction that perfectly fits in a measure of 4/4. Hence, whatever is played on the Disklavier will be played back exactly one beat after. Harmonically, with an axis on E4, F# and B played on the Disklavier will be played back as D and A on the next beat as seen in measure 23, which offer F#sus4 or D6/F# sound. If the left hand is shifted down a half step, with F and Bb played on the Disklavier and Eb and Bb in the playback, it will create F7sus4 or Bbsus4/F. Using these harmonic structures and the left hand patterns, the performer will play a solo which in effect turns into a duo. For instance, if a short motive such as one beat long motive, is played followed by a rest, the next moment during the rest is given to the mirror to play back that is rhythmically identical but harmonically inverted, which makes it sound somewhat different but somewhat similar. This also contributes to making it sound like two performers taking trades or two participants exchanging turns in a conversation making smooth transitions without overlapping.

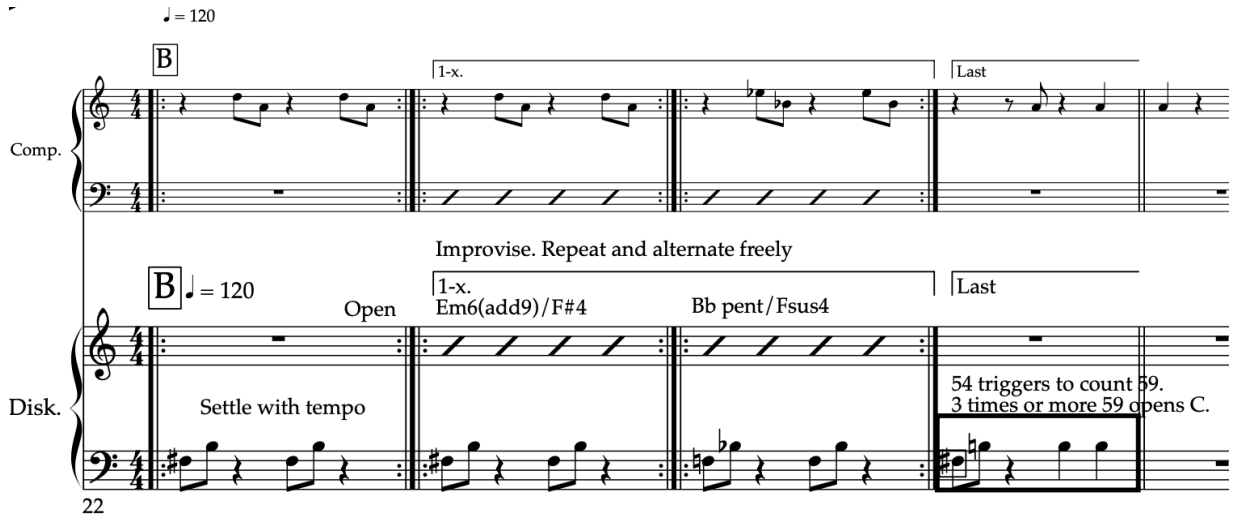


Figure 3.21 Section B, measure 22-25 of *Phasing*

Competitive overlap and Activity-occupied withdrawal

In section C, continuous shift of delay time is employed to depict gradual change of the display of engagement between the two participants in conversation, displaying activity-occupied withdrawal (Figure 3.22). Delay time which ramps from 0 to 900 milliseconds is applied over 20 seconds while the quintuplets in the right hand remain the same unaffected by the ramp by splitting the keys. This creates contrast between the right hand quintuplets that remain the same and the left hand which will be phased out over time. In the performance, the left hand of the performer begins with playing one single note at a time followed by a rest, giving a moment for the mirror to respond on the next beat. Then another note is added followed by a rest. After a regular interaction with a few words is established, the performer gradually adds more notes and plays continuous patterns of five or six notes which create the mirror to play back in exactly 500 milliseconds. This is to show that the conversation between the two initially started as a turn-taking with smooth transitions where both participants are present showing full

engagement. However, once the left hand starts playing continuous 5-note or 6-note figures with no rest in between, that causes the mirror to override non-stop as if the other participant decides to take over the original speaker. At this point, the status of this conversation has turned into continuous competitive overlap. This overlap will then gradually start to phase out as the delay time changes progressively over 20 seconds. With the delay time ramp from 0 to 900 milliseconds, after 20 seconds, the actual delay time applied changed from 500 to 1400 milliseconds including the built-in 500 milliseconds delay. In music at quarter note equals 120 bpm, 1400 milliseconds delay is a significant lag that is more than a glitch. Such a shift of delay time here, can thus manifest two participants moving away from each other from their initial engagement status in the conversation while still being present and occupied in the activity that is taking place, either because one is losing interest in the conversation or the level of discordance develops to a point where the two could no longer stay engaged.

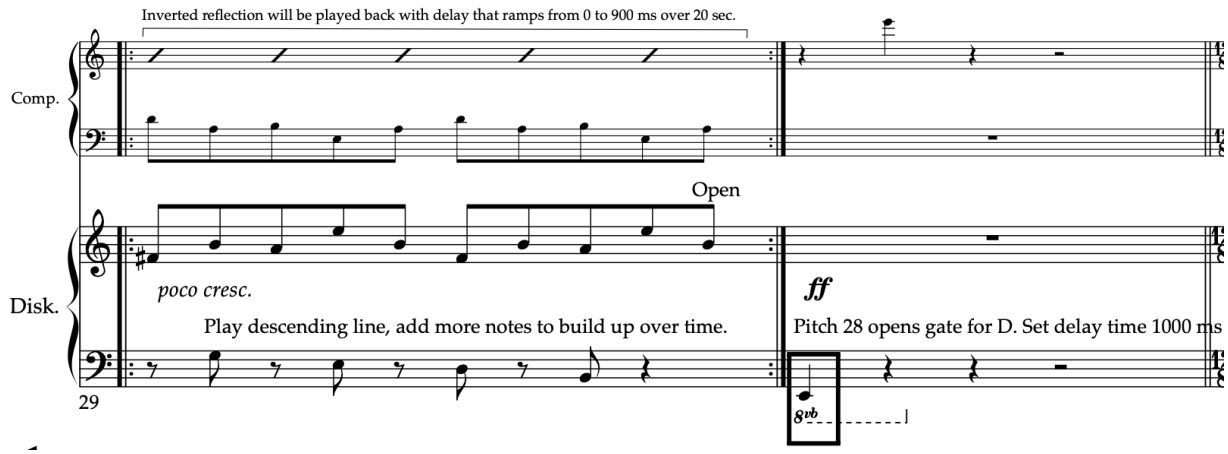


Figure 3.22 Measure 29-30 of *Phasing*

Turn-taking and mixed status of engagement

Section D is a representation of two participants taking turns in conversation in which they continuously take turns regardless of being overlapped with each other. In this section, the mirror axes change multiple times[17] which is triggered by the note 28 (E1) in the left hand. Every time E1 is played, the current axis changes to the next axis according to the index of the coll file in the max patch, which automatically changes the playback inversion while the input notes remain the same (Figure 3.23). For instance, from measure 37 to 38, the axis changes from E to G. Though the input notes remain the same as written, the playback will change from F# sus 4 to C sus4 as the axes change. In this way, the harmony progresses through eight different axes in this section during which various tensions, dissonances, pleasant or unpleasant harmonies are generated. These effects and impressions are to illustrate various stages of emotions and situations that may be experienced while being engaged in the dialogue in which they may be encountering conflicts, disagreements, arguments or moments of relief or consensus.

[17] These axes were chosen after experimenting with several different axes around E4 while always playing the same figure on the Disklavier. I learned that some of the axes cause the reflections to overlap with what I play on the Disklavier which is very troublesome during performance. I wrote out the reflections on the top staff and examined each combination, and arranged the order of axes to create gradual increase of dissonance in the outcome and adjust the range of the reflection. While the choice and the order of the axes are my own musical decision, it is also meant to reflect the changes in the displays of engagement in conversation.

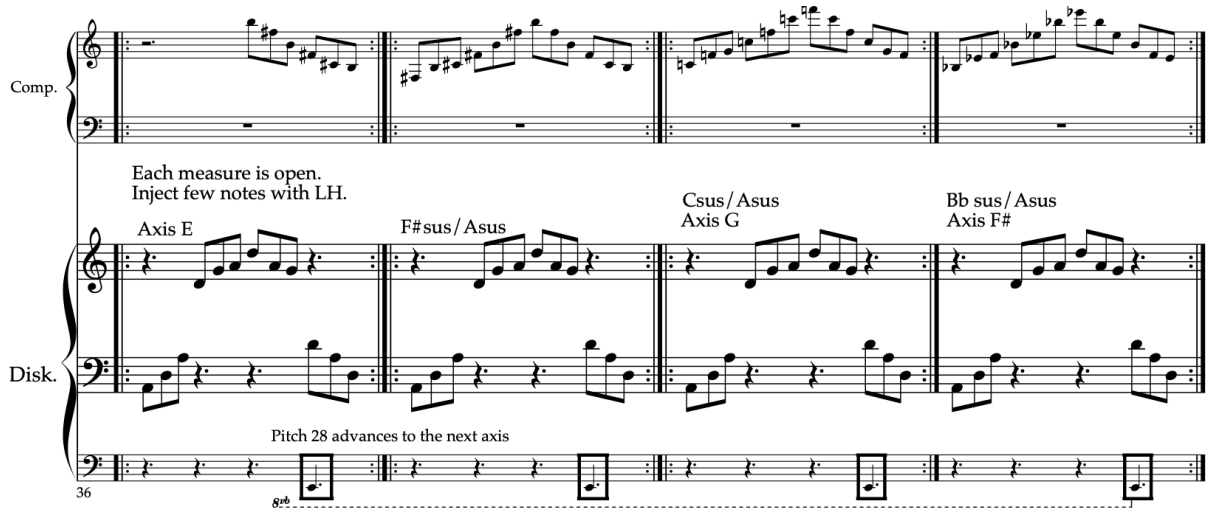


Figure 3.23 Section D, measure 36-38 of *Phasing*

Streams of Talk

Streams of Talk is an interactive solo piece for Disklavier programmed with MaxMSP. The music illustrates a conversational scene in which whenever a speaker initiates an utterance, there is always an interruption. An interruption in a form of whisper, shout or something in the middle which is released like a spray or a stream of bullets as if whenever one tries to say something, there is always someone who tries to talk over or cannot wait for his or her turn, giving you no chance to finish your sentence or even to start saying what you actually want to say. Another interpretation of this piece, though not necessarily conversational, could be a series of echoes which follows you around wherever you go and whatever you say. The original inspiration of using streams of notes comes from *Industrial* from *Natural Machines* by Dan Tepfer (Sunnyside, 2019).

The streams of notes are generated by a pitch offset that is dispersed when a key is depressed and released on the Disklavier. Any key depressed on the Disklavier will trigger either

an ascending or descending pitch offset that is transposed according to the pitch of the depressed key. The middle column shows the number of semitones from the depressed key for index 1 through 11, which are used as a pitch offset, and are dispersed when a note-off message of the depressed key is received at an interval of 120 milliseconds. Each pitch offset is assigned a velocity to create an increment of dynamics as shown in the rightmost column. Pitch F5 and anything below will trigger ascending lines after a note-off message of the depressed key is received. Any pitch above F5 will trigger descending lines.

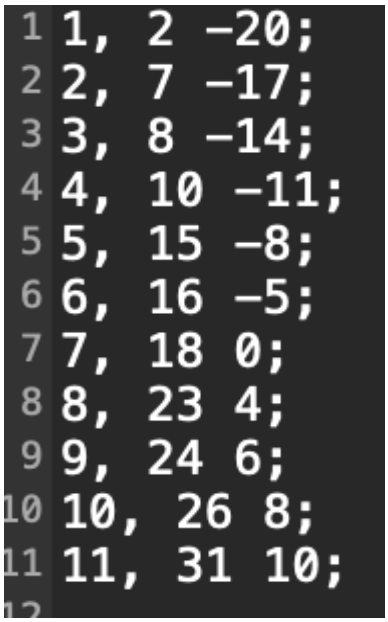


Figure 3.24 Ascending line

For instance, index 1 has a pitch offset 2 with a velocity -20. If the depressed key is a MIDI note number 60 with a velocity 40, it will generate a MIDI note 62 with a velocity 20 (Figure 3.24). Index 2 has a pitch offset 7 and a velocity -17, meaning it will generate a pitch that is five semitones higher than index 1 with a velocity that is slightly quieter than that of index 1. By doing so, each trigger pitch is designed to be transposed eleven times, always starting with a velocity 20 below incrementing up to 10 above the input note over 1200 milliseconds. Figure 3.25 shows an example of ascending pitches with a trigger pitch C1 (MIDI note 4). The numbers underneath the noteheads on the computer staff indicate the pitch offsets from the trigger pitch.

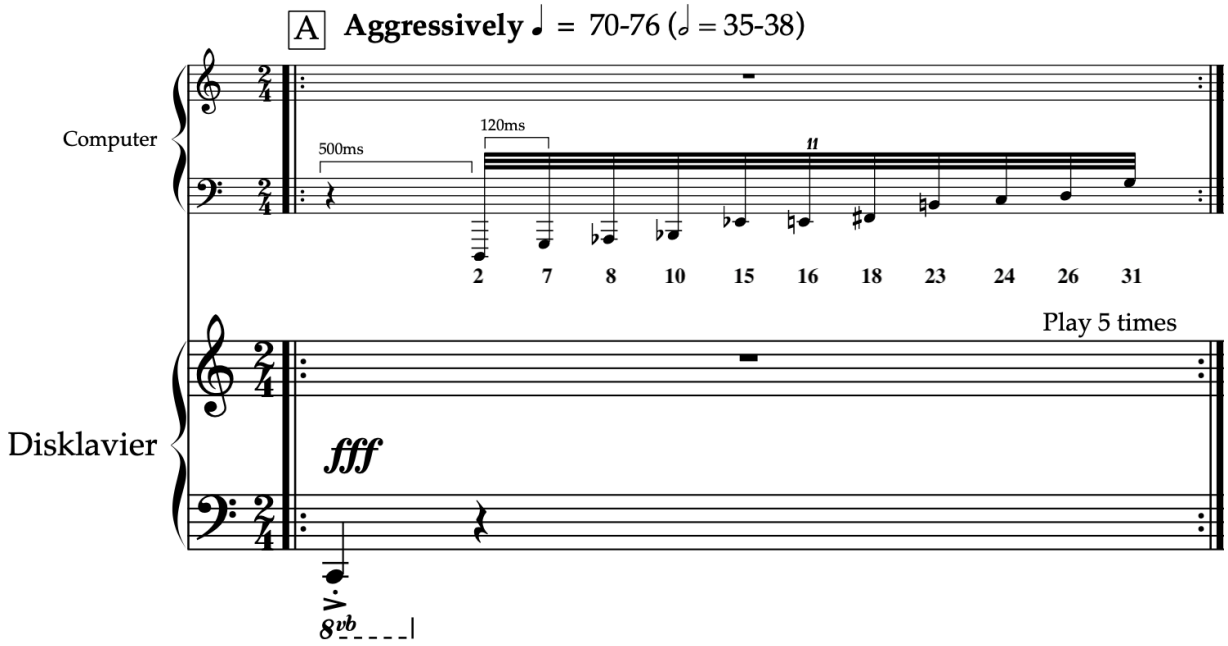


Figure 3.25 Ascending pitch offsets

```
1 1, -2 0;
2 2, -7 -3;
3 3, -8 -6;
4 4, -10 -9;
5 5, -15 -12;
6 6, -16 -15;
7 7, -18 -18;
8 8, -23 -21;
9 9, -24 -24;
10 10, -26 -27;
11 11, -31 -30;
```

Figure 3.26 Descending line

This pitch offset will be dispersed exactly 500 milliseconds after the note-off message of the input note (the depressed key) is received in delay mode in the Disklavier setting. Any input note below F# will create the transpositions of the pitch offset and increase in the velocity, creating an ascending series of 11 pitches. Descending lines which are the inversion of the ascending pitch offset will be generated when the input note is above F5 (Figure 3.26).

Turn-taking and Competitive pause overlap

At first, the speaker attempts to find a chance to take its turn by making short utterances, which is represented at the opening of the piece by the harsh articulation of C1 followed by a stream of 11 pitches that is repeated a number of times (Figure 3.27). Although the speaker is stopped after the initial utterance, the repetition of the beating and the stream that follows, takes the form of a turn-taking because the speaker waits until the completion of the streams, and they do not overlap. As regard to the type of human interruption, this would be classified as competitive pause overlap[18]. If that is the case, the original speaker may or may not attempt to interrupt themselves. Since the speaker does not attempt to take back his turn while the stream by the other party is being played here, we can confirm that this would be a type of competitive pause overlap. The built-in 500 milliseconds delay which happens every time creates an inevitable pause before the interruption. After a number of repeats, what follows after is a moment of pause during which the speaker realizes that every attempt will induce a stream of talk-back as shown in measure 2.

[18] Competitive pause overlap - Similar to "Collaborative pause overlap", a person might pause in conversation but before they can continue, the other person starts talking. In this case, the original speaker may or may not attempt to interrupt (Kane).

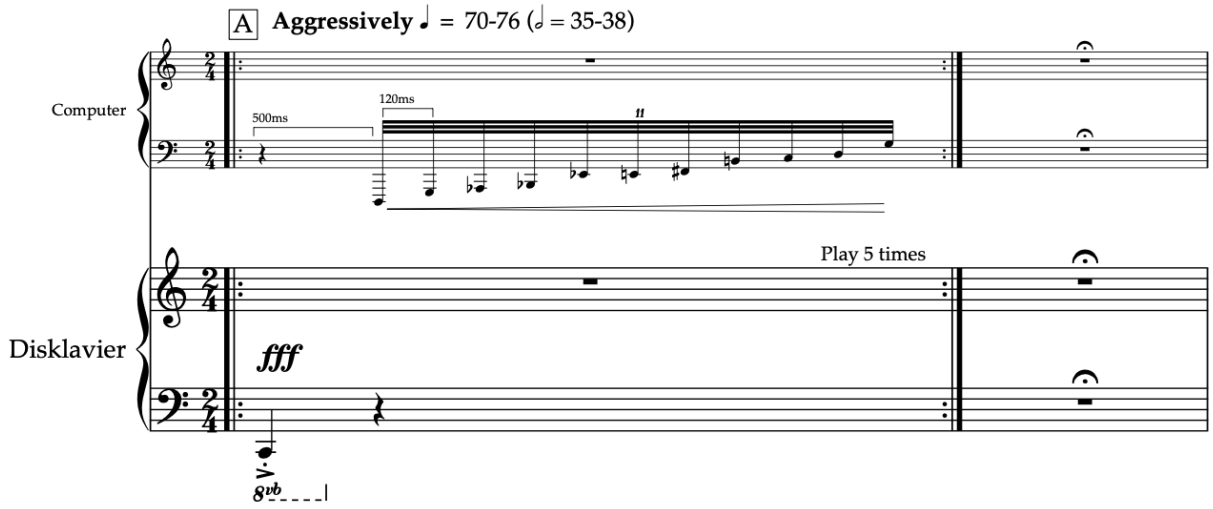


Figure 3.27 Opening of *Streams of Talk*

Competitive pause overlap and Speaking butting-in

Once realized, the speaker continues to challenge by beating notes steadily, and then gradually picks up the pace by making utterances before the completion of the streams as if the speaker chases back the other participant who keeps interrupting and talks over. This takes place in section B in which the tempo ramps up from 60 bpm to approximately 250 bpm over the next 30 seconds. Because the whole sequence takes exactly 1700 milliseconds from the moment a note off message of the pitch played on the Disklavier is received until the completion of the 11 pitches including the 500 milliseconds delay in between the receipt of the note-off message and the start of the stream, if the pacing of the beating accelerates, the stream and the next beating, or even the currently playing stream and the next stream will begin to overlap. Thus this is no longer a simple turn-taking where there is only one speaker at a time. Instead, this would be a competitive pause overlap which transitions into speaking butting-in.

Section C and D show multiple consecutive utterances being made by the speaker, either as fragments of motives made up of two to four notes, or two notes played at the same time. These sections were not necessarily performed with conversational scenes in mind, but primarily to depict two participants interacting with each other's responses regardless of the meanings or the outcome. Nevertheless, because any single note-off message will trigger either ascending or descending 11 pitches over 1700 milliseconds, these fragments of notes seen from measure 27 through 37 will result in doubling or tripling the streams, producing bullet-like notes (Figure 3.28). A possible interpretation of this situation can be that the speaker persistently attempts to initiate a conversation with a few words in hopes of saying a full sentence, but that only results in multiplying the amount of response, which overrides the speaker's sparse utterances.

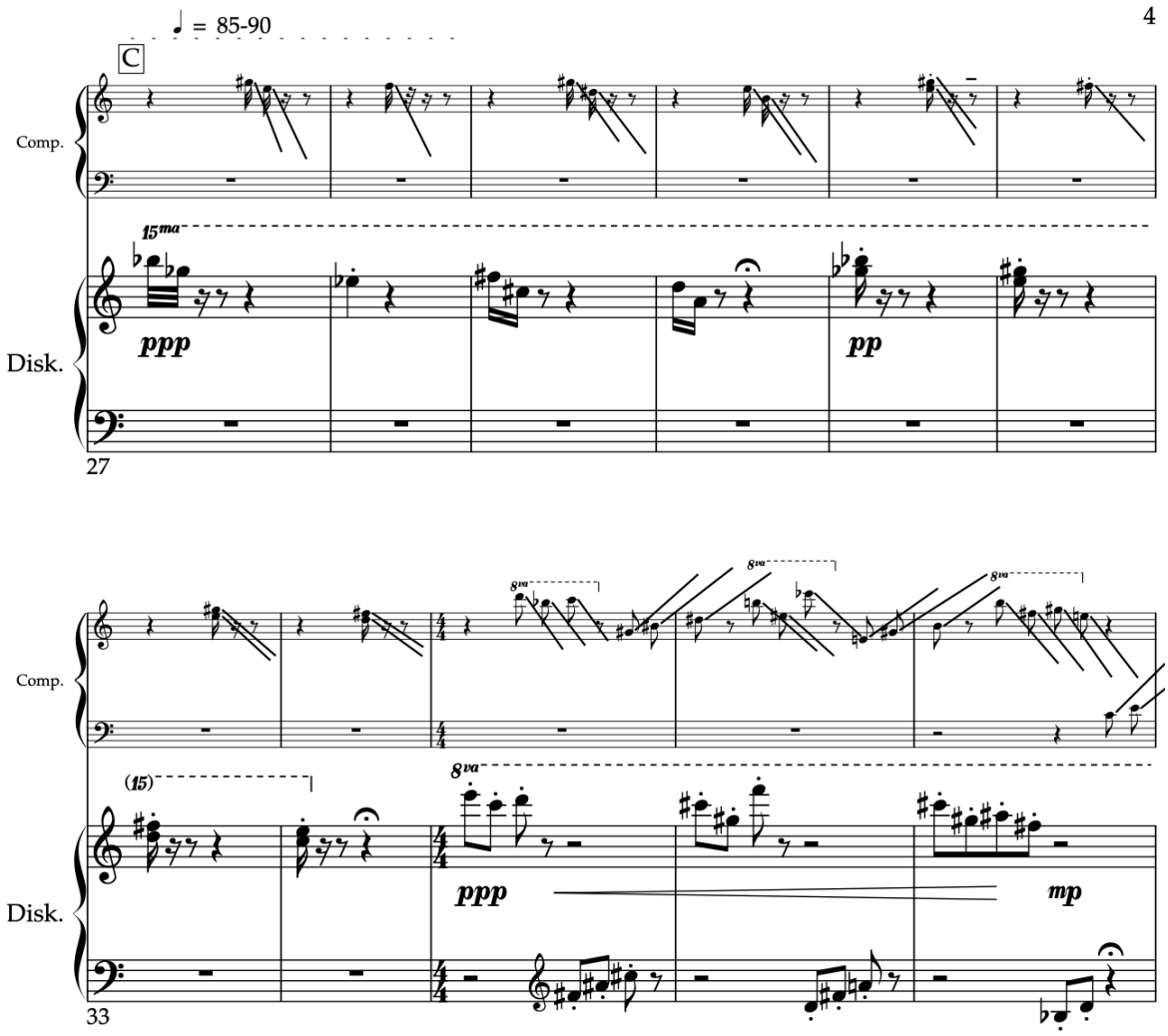


Figure 3.28 Measure 27-37 of *Streams of Talk*

If the speaker makes continuous utterances without giving any space for the other participant to be heard, it would create a cloud of streams in which the voice of the speaker and the other participant who makes the streams are no longer distinguishable, as seen towards the end of section D from measure 55, which is another form of speaking butting-in (Figure 3.29).

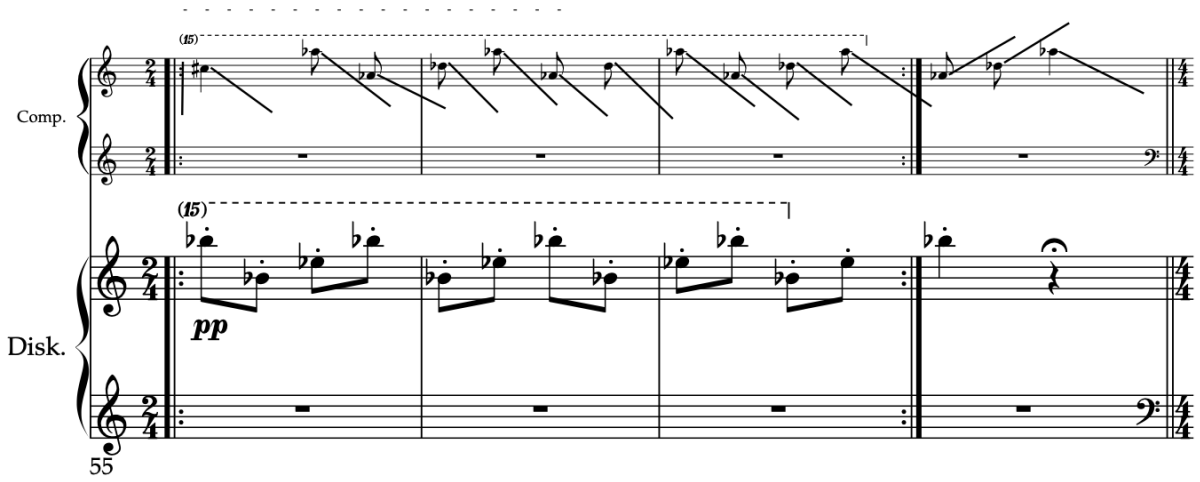


Figure 3.29 Measure 35, *Streams of Talk*

Turn-taking and Smooth transition

A slightly different attitude is seen in the recapitulation in section E, where it begins in the same way as the opening of the piece with a repetition of a single note beating harshly, which is beaten back immediately by the other participant in the form of streams as expected (Figure 3.30). However, this time, the speaker decides to multiply itself by stacking more notes in fifths and octaves, gradually expanding in size and sound. The intent is not only to amplify the sound, but also to say more in one short utterance with multiple voices stacked before the other participant tries to take over by its streams, which are also multiplied in accordance with the number of voices injected by the speaker. The outcome is a substantially more intense turn-taking with multiplied streams of responses, creating multi-layered smooth transitions. Although the speaker only has a fraction of a second given to take his or her turn before being interrupted by the stream of the other participant, this time, the speaker is making a strong statement, saying all he has to say in one single utterance by multiplying the voices before he gives the turn to the other participant. To make a statement, lengthy sentences may not be necessary. Sometimes, one

word or even one utterance of a sound such as uh, euh, shh, ah with a clear intention and a strong attitude can provide enough meaning and impact. The act of expanding the utterance by staking more voices can also illustrate one becoming more agitated or frustrated that he feels the urgency to elaborate his words every time.

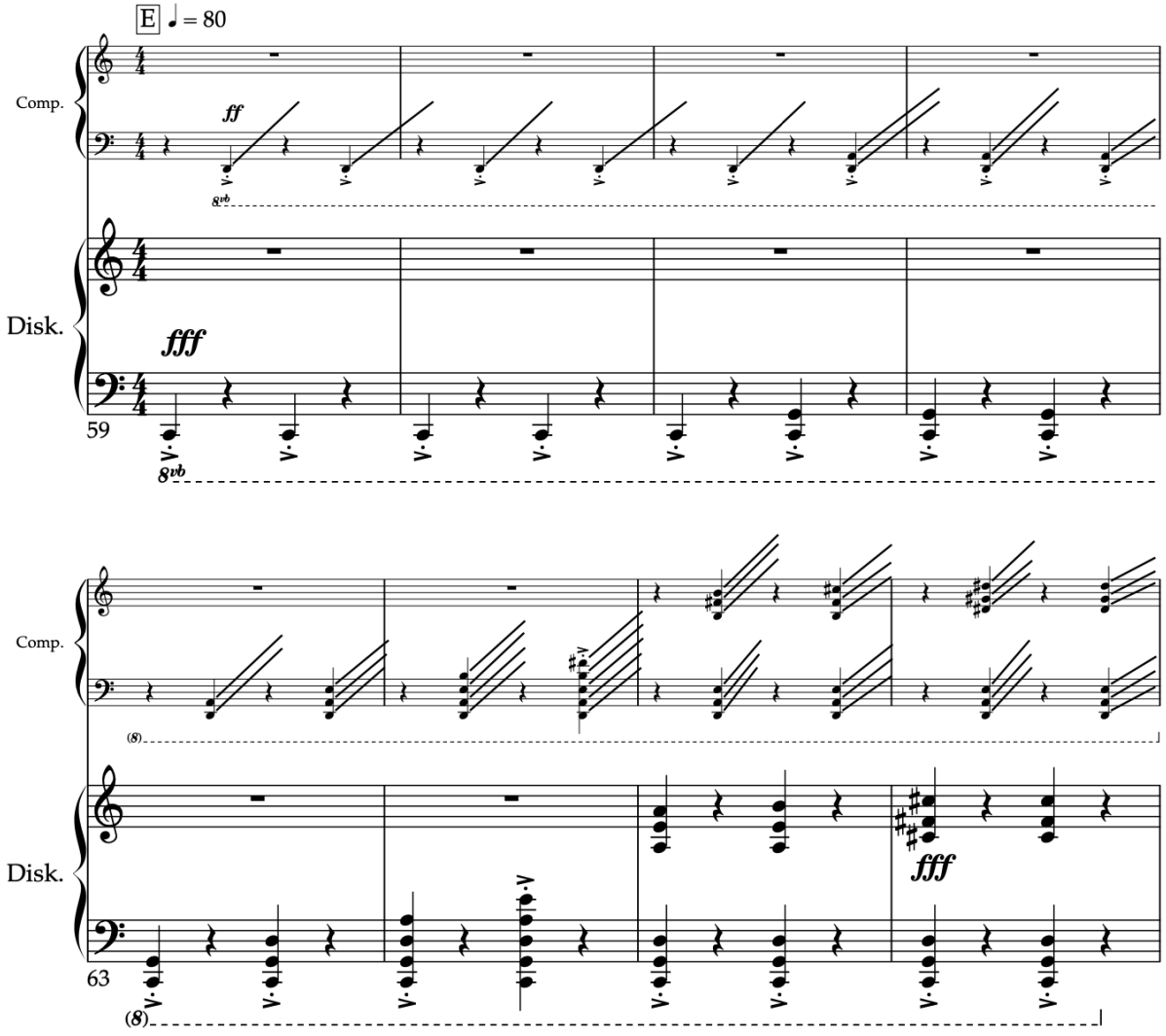


Figure 3.30 Measure 59-66 of *Streams of Talk*

Chapter 4: Evaluation

This chapter will assess the process of rehearsals and actual performances, compared to what was originally planned or assumed, such as things that worked out as expected, things that did not work so well, or things that I have learned after the performance.

Group Talk No. 1

In the actual performance, in section A which depicts turn-taking and accidental overlap, a fair amount of both turn-taking and accidental overlaps naturally occurred during the interaction between the piano and violin as they improvised. There were moments where one would play and pause to give space for the other to take the turn, and moments where both initiated their utterances simultaneously, but the overlaps occurred only briefly as the intent was not to take over the turn but to carry on to develop collaboratively.

The gestures of the violin and piano from measure 38 model mutual disengagement among two participants in conversation being occupied in different activities while being co-present in the same space. However, if this happened in actual conversation, it would most likely be an argument, as the two are speaking simultaneously but not listening to each other. For this to display a status of disengagement, one of them should be transitioning to other activities such as taking a cigarette gazing away from the speaker, or being distracted by TV while their lower body is still oriented towards the speaker showing some level of attentiveness. Having the two instruments play different musical materials simultaneously was intended to best illustrate a state of mutual disengagement. At the same time, it was a decision derived from musical needs.

The competitive overlap in section B, where all the players only make an utterance on beat 1, worked conceptually and as well as in the performance. The initial idea was to create accidental overlaps by interrupting the beginning of the utterance, which did happen during the performance by injecting disturbing sounds when somebody initiates an utterance on beat 1. However, in other times, it turned out to be more like a call and response or turn-taking where somebody plays on beat one and other people make a sound that is disturbing or completely unrelated during the space after beat 1. Because it did not affect the flow of the music, I believe that was what the music called for, and was how the performers interpreted it while playing.

Group Talk No. 2

The opening of *Group Talk No. 2*, where the string players were assigned specific pitches and beats, allowed smooth transition in turn-taking. Despite the constrictions placed on the performers in terms of the placement of the sound and the pitch, I consider this smooth transition was made possible without interruption because of the constrictions and the framework within which they were allowed to embellish or orchestrate. The injection of external interruption by the cues from the percussion was not only a musical addition, but also comes from other possible situations such as phone calls or siren from outside, that were not mentioned in the principles of conversation I have researched.

The operations by recipient and backchanneling demonstrated in section C worked as expected and notated. The only unusual behavior, were it an actual conversation, is that the backchanneling by the hearer begins prior to the utterance by the speaker, which is impossible. The reason why I adapted this way is because the recurring 8-bar phrase in the following section,

section D, always begins two bars ahead. Since both sections are based on the same 8-bar phrase with the same ostinato and harmony, I chose to keep the same format starting section C.

The turn-taking in section D worked as expected. The string players were able to trade their turns by creating new rhythms each time, and were backchanneling on the upbeat of three to show attentiveness to the speaker. Accidental overlap and collaborative pause overlap also happened during the performance. The only moment that was unexpected is when both players suddenly seemed to behave hesitantly after they performed one cycle of overlap which is seen right after the third iteration when neither of them initiated for the next 8-bar cycle (Figure 4.1). While this appears to be an accidental issue in performance, this could also be a common situation to follow after a moment of overlap in actual conversation that both parties hesitate for a moment to figure out who will take the next turn, and end up choosing to remain silent.

Another point to note about this section is the practicality of collaborative pause overlap in conversation. According to Dr. Kane "the original speaker lets the other person continue, making the interaction sound collaborative or neutral" (Kane, 2) after the second person starts speaking, which is perfectly depicted in the third iteration and works fine musically. My question is whether such an overlap can still maintain a neutral and collaborative manner when two people are speaking simultaneously. On the other hand, what happened during the third iteration starting from 3:13 is the two different rhythms interlocking with each other and resolving with a suspension in the fifth measure. If two participants of conversation continued to talk simultaneously, it would not have sounded organized or collaborative. Thus what may sound competitive or not collaborative in actual conversation can be converted into a collaborative outcome when translated into music.

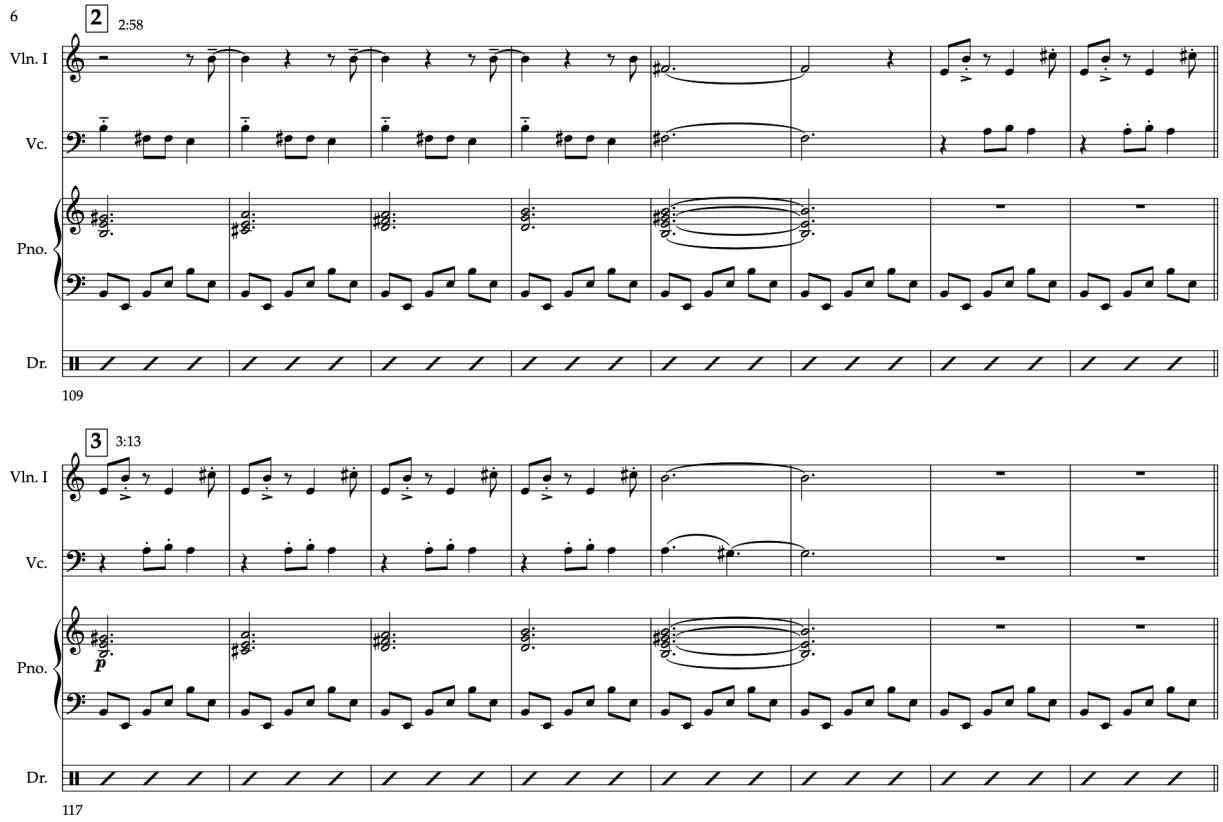


Figure 4.1 Section D transcription of *Group Talk No. 2*

Section E, where gradual transition of participants entering the state of engagement and disengagement is depicted, has yielded a musical outcome that was not originally planned but worked. The entering and transitioning to stages of engagement and disengagement were achieved successfully by the performers joining with their rhythms one by one and modifying them individually as well as developing as a group. This resulted in interlocking multiple lines of rhythms (Figure 4.2). In some cases, for instance the cello, began to extend motives by adding notes and displacing the entrance, which created phrases that cross over the bar lines. Musically speaking, the effect of having interweaving rhythm certainly contributed to making the music develop more collaboratively. Although this type of interaction may not happen in spoken

conversation, it does model two participants overlapping when they are in agreement. It was the concept of engagement and disengagement, while being in the same space and still being occupied in the same activity, that offered the structure to develop this section.

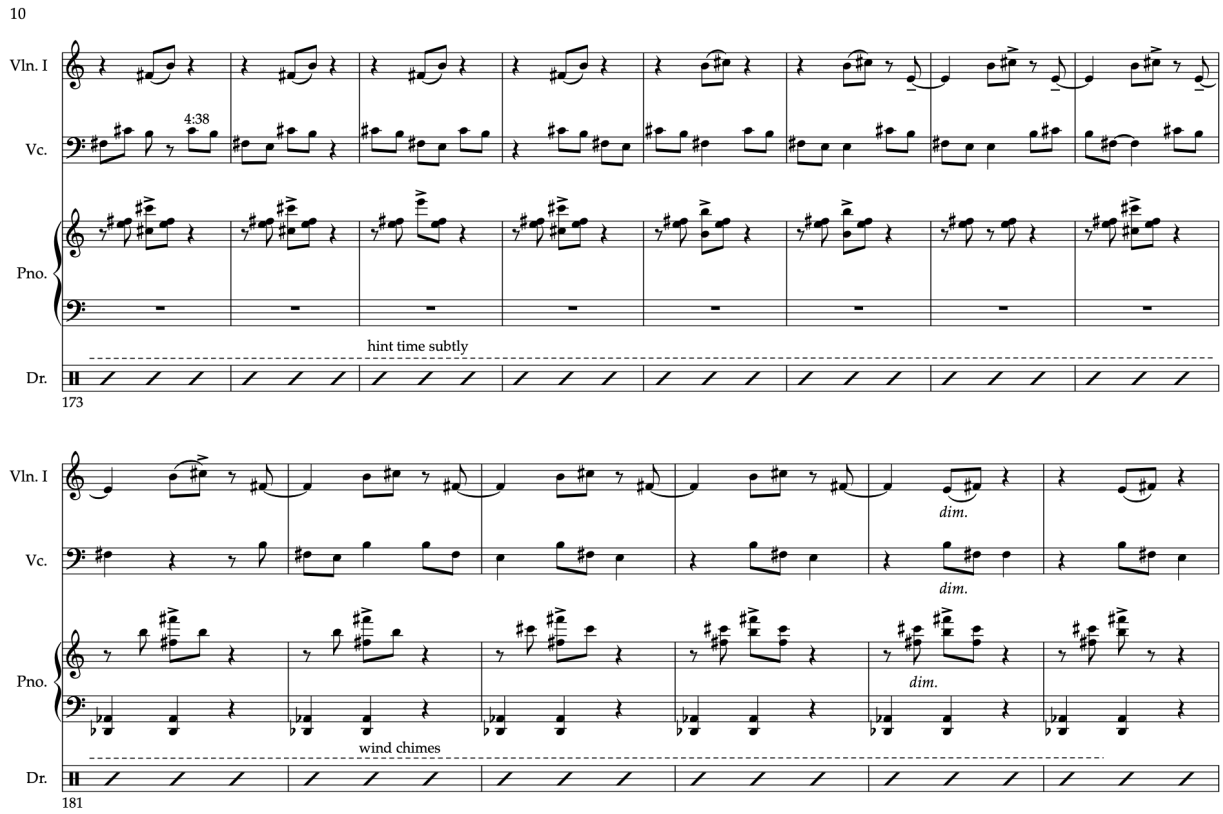


Figure 4.2 Section E transcription of *Group Talk No. 2*

Group Talk No. 3

In the performance of section A, the abstract image of collaborative overlap was manifested and achieved by the four performers creating swells at various rates and intensity at their own pace while maintaining mutual accommodation in the interaction. There were moments where swells arose one by one, and moments where two would emerge together, and another

would rise as the two began to fade. These gestures can be considered an interpretation of a group talk in which participants freely initiate talks and take turns in non-orderly ways yet maintain a collaborative and supportive attitude.

Section B of Group Talk No. 3 explored types of withdrawals by speakers and hearers as the performers fluidly moved from a shape to another shape. The five shapes given to the performers are labeled as boxed numbers in the transcription below (Figure 4.3). We can see right from the beginning that the shape 1 initiated by the violin is immediately joined by the cello, but the violin withdraws and engages in another shape 2 from around 2:24 mark while the cello was still playing shape 1. Shape 2 was originally initiated by the piano before the cello joined, hence this shows that the cello was noticing what the piano was playing. The quintuplets which the cello begins from 2:40 is joined by the piano after five seconds, but the cello slowly shifts away from the shape to transform into a slow version of shape 5 which is as if the cello started engaging in other activity disengaging from the piano while still partially participating in the ongoing talk with the piano.

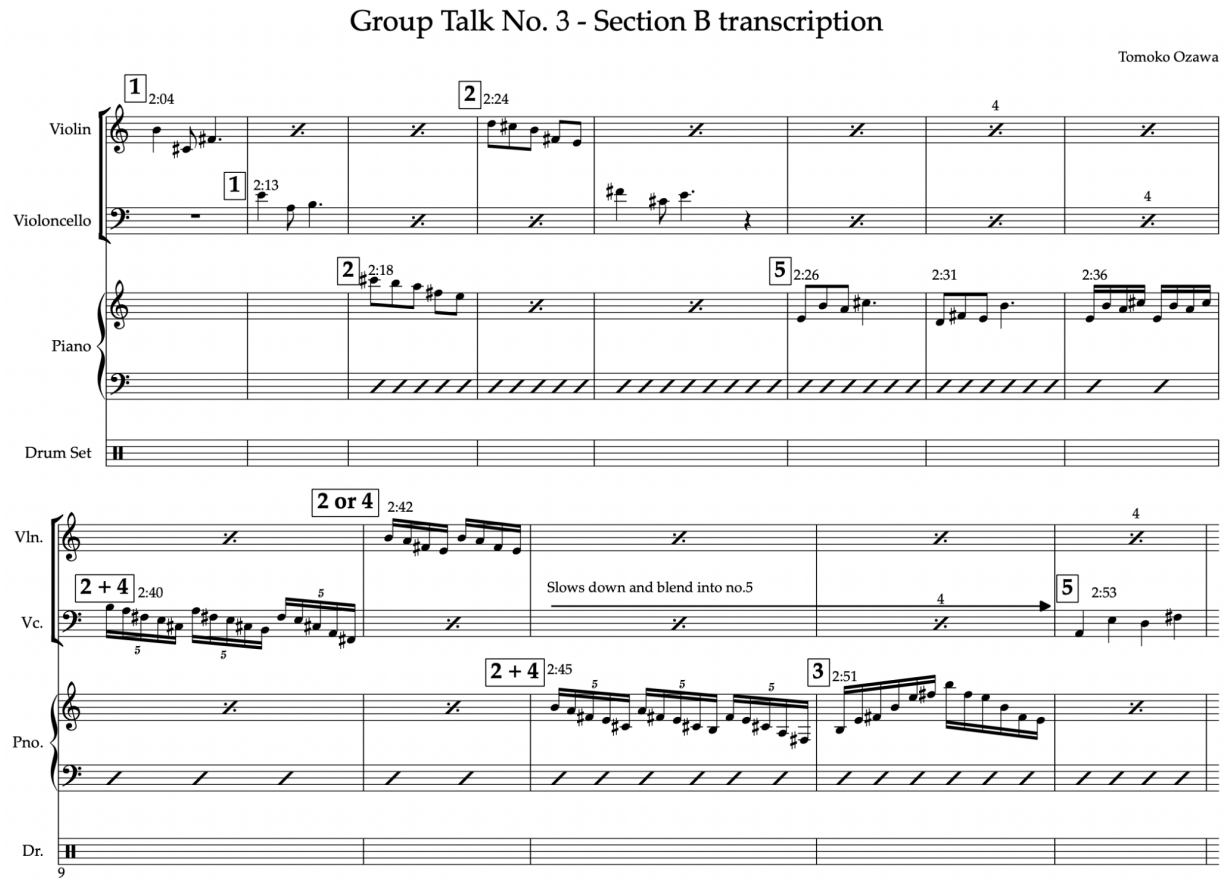


Figure 4.3 Section B transcription of *Group Talk No. 3*

What is observed after the cue was first called by the violin at 2:57 is gestures of refusing to withdraw (Figure 4.4). As soon as the violin cued, it was noticeable to all other players. The fact that cello and piano joined after few repetitions of the cue were heard shows them refusing to withdraw from the shapes they were playing at that time. We can see that the concepts of engaging and withdrawing from shapes were recreated as part of performance.

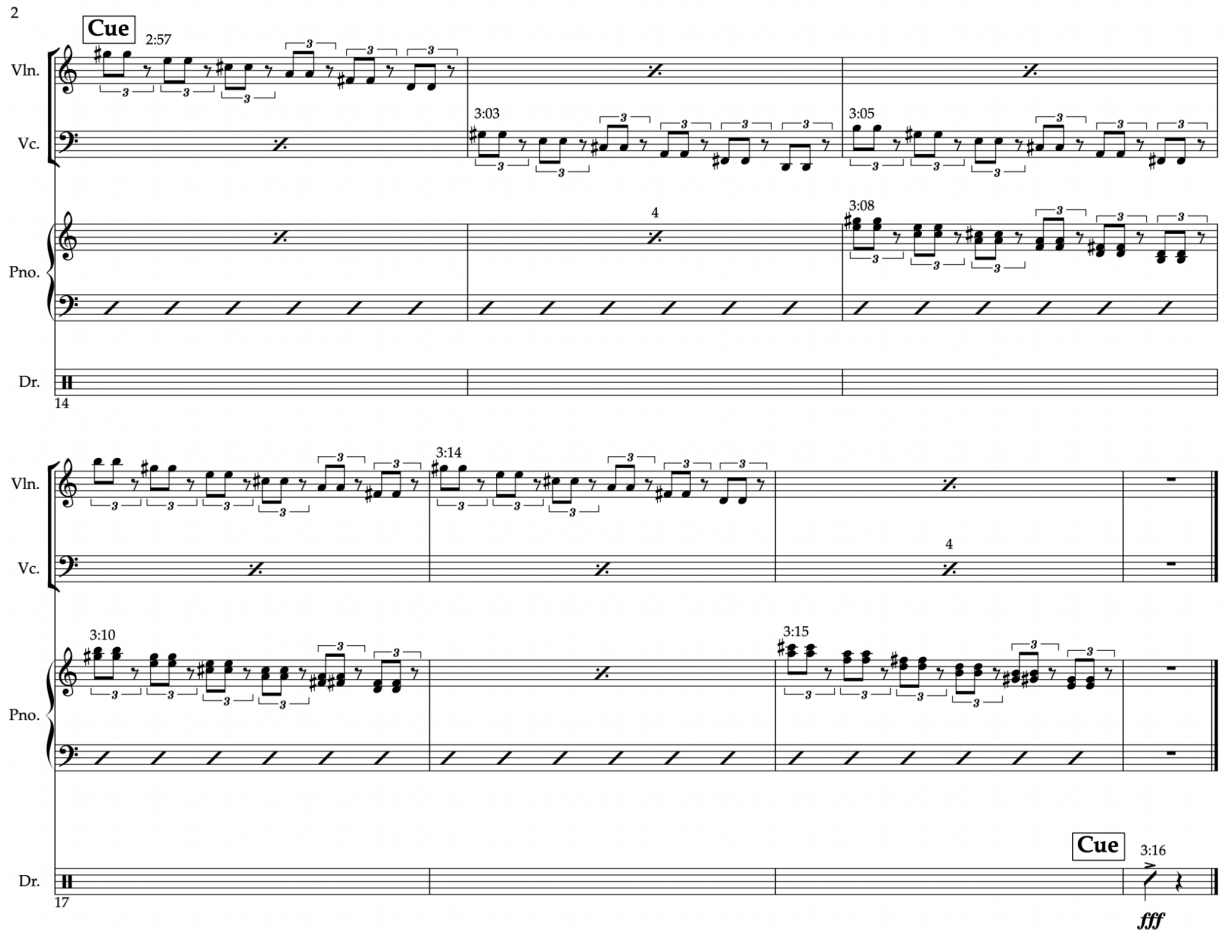


Figure 4.4 Section B cues, *Group Talk No. 3*

Cricket Wind

Smooth transitions and operations by the recipient were performed to recreate a dialogue between the piano and the wind created by the computer. Though the wind was always initiated manually by the vertical movement of the right hand, having the idea of turn-taking between the piano and computer has given the structure to formulate the phrasing. At the same time, because the vertical gesture of the wrist occurs from playing the piano, the interaction between the piano

and computer was invoked in an organic way to make smooth transitions and reactions from the computer.

In a similar yet opposite way, smooth transition and operations by the recipient were achieved in section B, where the sound of piano was manipulated and delayed before being played back, triggered by the inactivity of the hand. Because the response always relied on the input from the piano and the timing until the hand came to stillness, the response of the computer was unique each time, which enabled turn-taking to be smooth transition without overlap and allowed the computer to produce various responses.

Phasing

In this piece, a number of conversational models were explored such as speaking butting-in (in the opening), turn-taking (in section B), competitive overlap and activity-occupied withdrawal, and mixed status of engagement. The feature that contributed the most to making a dialogue in a solo piece is the reflection created by inversion using axes. The reflection enabled any phrases fed into Disklavier to be played back exactly in its inverted form, hence played on the opposite side of the keyboard. By being able to adjust the timing of playback from 500 milliseconds or more, the response to the playing was programmable at a fixed timing or using continuously growing delays. Fixed delays were used to recreate speaking butting-in in the opening and for the turn-taking in section B and D. Continuously growing delays were used in section C to depict gradual change of the display of engagement between the two participants in conversation, displaying activity-occupied withdrawal. A combination of a fixed delay and

changing axes contributed to recreating various stages of engagement and disengagement while the playing and the reflection experienced phases of overlaps through eight different harmonies.

Streams of Talk

In this piece, despite having only one algorithm with the 11 pitch offsets, various situations of conversation such as turn-taking, speaking butting-in, competitive pause overlap, and smooth transition, were manifested in the improvised performance. Although this piece was 100% improvised, ideas for each section and the structure were constructed prior to the performance. Using the 11 pitch offset that are always dispersed at fixed time intervals over 1700 milliseconds, each situation of conversation was reimagined and recreated using different dynamics, registers, intervals and timing or tempo.

For instance, the opening of the piece uses constant beating of C1 to trigger the 11-pitch stream to recreate turn-taking between the Disklavier and the stream, while using the 500 milliseconds delay to exemplify competitive pause overlap which prevents the Disklavier from continuing to talk. In contrast, section B uses accelerating tempo to create multiple incidents of competitive pause overlap and speaking butting-in. Section C and D use consecutive utterances to depict the speaker attempting to say more than one utterance though it only induces more overlap to follow. As the speaker accelerates the speed, the conversation moves towards the highest register while diminishing the sound. This blurs the distinction between the speaker (the Disklavier) and the stream, creating continuous speaking butting-in that is heard from afar. The recapitulation in section E uses the beating in the lowest register playing with fortississimo to amplify the impact of the speaker's utterances as well as to depict the speaker attempting to

maximize each utterance by stacking the notes though it results in multiplying the interrupting response.

Despite having one single gesture of 11 pitches as the foundation of the piece, and the only possible variation being the direction of the stream, multiple situations of conversation were illustrated throughout the piece in the improvised performance. While in actual conversation, these bullets of talkbacks will unlikely occur to this degree, it is the models of conversation that provided conceptual structures to the music. Use of these models have enabled me to translate various conversational scenes into music to construct structures, and navigate the music through the imaginative scenes.

What I learned through the process of composing music using models of conversation, despite the tensions I encountered between the music and the concept I intended to apply, is that it offered a different compositional approach. Instead of developing compositional ideas based on traditional theories of harmony or techniques of developing motives, the music could be created using the structures, scenarios or situations taken from conversational scenes which function as frameworks. In addition to creating frameworks, transitions between scenes could also be drawn from conversation and be translated into music.

Chapter 5: Summary and conclusions

I have explored the use of various conversational models as composition techniques in group conversation settings in *Group Talk No. 1-3* and as a dialogue in my solo pieces *Cricket Wind, Phasing,* and *Streams of Talk*. Instead of using the concept of conversation in a metaphoric way after the music is composed or performed, the intention of my research was to use specific models of conversation to construct the activities in the music.

While each phase of conversation, whether it be turn-taking or a stage of interruption, could be translated into music by my own interpretation, the challenge I encountered at times was the tension between the music and the concept. In some cases, it could be adapted into music but it might not be a common practice in actual conversation. For instance, overlaps do occur commonly in any conversation or dialogue for a brief period of time. When these happen, participants will quickly act to correct it. Usually the participant will try to reorganize to have one speaker at a time instead of allowing two or more people to continue to speak simultaneously. However, if such a situation occurs in music, there is more than one possibility to maintain simultaneous speech. The composer has the ability to create the kind of relationship during the overlap by making the players play in the same tonality to create a collaborative spirit or make them play in different tonalities or meters or at different speeds to create cacophony. In other cases, the music calls for its own directions that are not conversation-oriented, but inspired by other instances or derived from the traditional composition techniques. When I encountered those moments while writing these compositions based on models of conversation, I chose to allow both conversational devices and other inspirations to blend and navigate the music.

Because it is a conceptual structure which conversational composition provides, it does not conflict with use of other composition techniques. Rather, it can be used in combination with other composition techniques to offer different ways of structuring the music which is not possible in the traditional composition methods. In addition, the use of conversational models in composition can shed further light and understanding on what is already happening in improvised music, hence offering the composer more ways to organize interactions in the music.

I have learned two lessons from composing using these models of conversation. First, employing these compositional devices has led me to compose in new ways; the music grew explicitly from an extra-musical source, providing a conceptual structure for the performers' interactions. Use of conversational models such as ordering of turn-taking, various types of overlaps, ways to maintain conversation through different phases of engagement or disengagement, as compositional devices provides structural framework to guide through the music in a way that is distinct from composing in the traditional composition methods. Second, because the music itself represents conversation, it organically leads the performers to engage more actively with each other in being part of the conversation as a collective. During the rehearsals and concert, the performers are made more aware of their surroundings compared to conventional jazz settings where the script of ordering or events is planned in advance. It is the framework of composition derived from conversation that grants agency to the performers, allowing them to express themselves while contributing to the composing process as the music is driven by conversation. Actively employing these techniques of conversational models results in a highly interactive performance.

Bibliography

Angelino, Lucia. “Collective Intentionality and the Further Challenge of Collective Free Improvisation.” *Continental Philosophy Review* 53, no. 1 (2020): 49–65. https://doi.org/10.1007/s11007-020-09484-y.

Chomsky, Noam. *Syntactic Structure*. The Hague: Mouton, 1957.

Chomsky, Noam. *Aspects of the Theory of Syntax*. Cambridge, MA: The MIT Press, 1965.

Clayman, Steven E., and Tanya Romaniuk. “Chapter 1. Questioning Candidates.” *Talking Politics in Broadcast Media*, 2011, 15–32. https://doi.org/10.1075/dapsac.42.05cla.

Clayman, Steven E., and John Heritage. “Questioning Presidents: Journalistic Deference and Adversarialness in the Press Conferences of U.S. Presidents Eisenhower and Reagan.” *Journal of Communication* 52, no. 4 (2002): 749–75. https://doi.org/10.1111/j.1460-2466.2002.tb02572.x.

Crystal, David. *Let's Talk: How English Conversation Works*. Oxford: Oxford University Press, 2020.

Coleman, Ornette. Liner notes for Coleman, Ornette. *The Change of the Century*. Performed by Don Cherry, Charlie Haden, Billy Higgins and Ornette Coleman. Recorded 1959. New York: Atlantic, 1960. CD.

Coleman, Ornette. *Free Jazz Ornette Coleman*. CD. A&R Studios, NY: Atlantic, 1960.

Drew, Paul, and John Heritage. *Conversation Analysis*. London: Sage Publications Ltd, 2006.

Drew, Paul, and John Heritage. *Contemporary Studies in Conversation Analysis*. London: SAGE, 2013.

Freud, Sigmund, James Strachey, and Angela Richards. *Introductory Lectures on Psychoanalysis*. London: Penguin, 1991.

Garfinkel, Harold. “On the Origins of the Term ‘Ethnomethodology.’” Essay. In *Ethnomethodology Selected Readings*, edited by Roy Turner, 15–18. Harmondsworth: Penguin Education, 1974.

Goodwin, Charles. *Conversational Organization: Interaction between Speakers and Hearers*. Academic Press, 1981.

Goffman, Erving. “Alienation from Interaction.” *Human Relations* 10, no. 1 (1957): 47–60. https://doi.org/10.1177/001872675701000103.

Goffman, Erving. “The Interaction Order: American Sociological Association, 1982 Presidential Address.” *American Sociological Review* 48, no. 1 (1983): 1. https://doi.org/10.2307/2095141.

Griffith, Jennifer. “Mingus in the Workshop: Leading the Improvisation from New Orleans To Pentecostal Trance.” *Black Music Research Journal* 35, no. 1 (2015): 71–95. https://doi.org/10.5406/blacmusiresej.35.1.0071.

Gumperz, John J., and Dell Hymes. *The Ethnography of Communication*. 66. Vol. 66. 2. Washington, D.C.: American Anthropological Association, 1964.

Hagberg, Garry. “Ensemble Improvisation, Collective Intention, and Group Attention.” Essay. In *The Oxford Handbook of Critical Improvisation Studies* 1, edited by George Lewis and Benjamin Piekut, 1:481–99. New York, NY: Oxford University Press, 2016.

Holt, Elizabeth. “The Last Laugh: Shared Laughter and Topic Termination.” *Journal of Pragmatics* 42, no. 6 (2010): 1513–25. https://doi.org/10.1016/j.pragma.2010.01.011.

Jefferson, Gail. “Notes on ?Latency? in Overlap Onset.” *Human Studies* 9, no. 2-3 (1986): 153–83. https://doi.org/10.1007/bf00148125.

Kane, John. “Nine Types of Human Interruptions in Conversations.” *Cogito*, 3 Feb. 2020, www.cogitocorp.com/researchers-identify-nine-types-of-human-interruptions-in-conversations/.

Kendrick, Kobin H., and Francisco Torreira. “The Timing and Construction of Preference: A Quantitative Study.” *Discourse Processes* 52, no. 4 (2015): 255–89. https://doi.org/10.1080/0163853x.2014.955997.

Kuroshima, Satomi. “Another Look at the Service Encounter: Progressivity, Intersubjectivity, and Trust in a Japanese Sushi Restaurant.” *Journal of Pragmatics* 42, no. 3 (2010): 856–69. https://doi.org/10.1016/j.pragma.2009.08.009.

Kushida, Shuya. “Confirming Understanding and Acknowledging Assistance: Managing Trouble Responsibility in Response to Understanding Check in Japanese Talk-in-Interaction.” *Journal of Pragmatics* 43, no. 11 (2011): 2716–39. https://doi.org/10.1016/j.pragma.2011.04.011.

Monson, Ingrid. *Saying Something: Jazz Improvisation and Interaction*. University of Chicago Press, 1997.

Morris, Lawrence Butch, Daniela Veronesi, Allan Graubard, J. A. Deane, Massimo Golfieri, Luciano Rossetti, and Howard Mandel. *The Art of Conduction: A Conduction Workbook*. Edited by Daniela Veronesi, Allan Graubard, and J A Deane. New York, NY: Karma, 2017.

Pillay, Rona. "Ethnomethodology." *Handbook of Research Methods in Health Social Sciences*, 2019, 269–83. https://doi.org/10.1007/978-981-10-5251-4_68.

Pillet-Shore, Danielle. "Doing Introductions: The Work Involved in Meeting Someone New." *Communication Monographs* 78, no. 1 (2011): 73–95. https://doi.org/10.1080/03637751.2010.542767.

Tepfer, Dan. "Dan Tepfer's Natural Machines - Full Album." YouTube. YouTube, October 27, 2018. https://www.youtube.com/playlist?list=PLUzNnspxG1f5Bks6R8ckCzDSEn__psaO8.

Tepfer, Dan. "Natural Machines." Dan Tepfer, October 27, 2018. https://dantepfer.com/blog/?p=711.

Raevaara, Liisa. "Accounts at Convenience Stores: Doing Dispreference and Small Talk." *Journal of Pragmatics* 43, no. 2 (2011): 556–71. https://doi.org/10.1016/j.pragma.2010.01.020.

Risset, Jean-Claude. Instructions For Playing This Piece. *Duet for One Pianist.* Massachusetts: Massachusetts Council of the Arts and Yamaha France Musique, 1989.

Risset, Jean-Claude. "From Piano to Computer to Piano." In *ICMC Glasgow 1990: Proceedings*, 15–19. Glasgow: International Computer Music Association, 1990.

Risset, Jean-Claude, and Scott Van Duyne. "Real-Time Performance Interaction with a Computer-Controlled Acoustic Piano." *Computer Music Journal* 20, no. 1 (1996): 62–75. https://doi.org/10.2307/3681273.

Robinson, Jeffrey D. "An Interactional Structure of Medical Activities during Acute Visits and Its Implications for Patients' Participation." *Health Communication* 15, no. 1 (2003): 27–59. https://doi.org/10.1207/s15327027hc1501_2.

Sacks, Harvey. Essay. In *Studies in Social Interaction*, 280–93. Notes on Police Assessment of Moral Character. New York: Free press, 1972.

Sacks, Harvey, Emanuel A. Schegloff, and Gail Jefferson. "A Simplest Systematics for the Organization of Turn-Taking for Conversation." *Language* 50, no. 4 (1974): 696. https://doi.org/10.2307/412243.

Sacks, Harvey, Gail Jefferson, and John Heritage. “Notes on Methodology.” Essay. In *Structures of Social Action Studies in Conversation Analysis*, edited by J. Maxwell Atkinson, 21–27. Cambridge Cambridgeshire: Cambridge University Press, 1984.

Sacks, Harvey. “The M.I.R. Membership Categorization Device.” *Harvey Sacks Lectures 1964–1965* 12 (December 1989): 271–81. https://doi.org/jstor.org/stable/20009063.

Sacks, Harvey, and Gail Jefferson. *Lectures on Conversation*. 2. Vol. 2. Oxford: Blackwell, 1995.

Sauer, Theresa. *Notations 21*. New York, NY: Mark Batty, 2008.

Schegloff, Emanuel A. 1963. “Toward a Reading of Psychiatric.” Theory. *Berkeley Journal of Sociology.* California: University of California. Vol. 2: 61-91. http://jstor.org/stable/42809111.

Schegloff, Emanuel A. 1968. “Sequencing in Conversational Openings.” *American Anthropologist*. https://doi.org/10.1525/aa.1968.70.6.02a00030.

Schegloff, Emanuel A. “A Tutorial on Membership Categorization.” *Journal of Pragmatics* 39, no. 3 (2007): 462–82. https://doi.org/10.1016/j.pragma.2006.07.007.

Schuller, Gunther. *Early jazz: Its roots and musical development*. New York: Oxford University Press. 1968.

Sidnell, Jack. 2011. *Conversation Analysis: An Introduction*. John Wiley & Sons.

Sidnell, Jack, and Tanya Stivers. *The Handbook of Conversation Analysis*. Malden, MA: Wiley Blackwell, 2014.

Stivers, Tanya, and Federico Rossano. “Mobilizing Response.” *Research on Language & Social Interaction* 43, no. 1 (2010): 3–31. https://doi.org/10.1080/08351810903471258.

Stokoe, Elizabeth. *Talk: The Science of Conversation*. Little, Brown Book Group Limited, 2021.

Walker, William Franklin. *A Conversation-based Framework for Musical Improvisation,* PhD Thesis, University of Illinois at Urbana-Champaign, 1994.

Appendix

Scores, transcriptions and videos

http://www.tomoko-ozawa.com/phd