Communication and Coordination Between Singers Performing Duets

by

Ella Fund-Reznicek

B.A., Linguistics, University of Kansas, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Arts

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Linguistics)

The University of British Columbia
(Vancouver)

March 2016

Abstract

How do singers communicate with each other while performing? In this study, pairs of singers rehearsed three duets separately, and then performed them together for the first time in the lab. Since singers cannot talk to one another during a performance to assist in coordinating their actions, they must use other modes of communication to synchronize their performances. Correlation map analysis and transfer entropy analysis of video and audio recordings of these performances examined how the singers coordinated their performances non-verbally over the course of each recording session.

Preface

This thesis is the original, unpublished work of the author, E. Fund-Reznicek. All data was collected under UBC Ethics Certificate #B04-0558 (”Determining Event Structure in Communicative Interaction”; Principal Investigator, Eric A. Vatikiotis-Bateson).

Table of Contents

List of Tables

List of Figures

Acknowledgments

First, I wish to thank my committee members for their ideas, guidance, and support. Specifically, I'd like to thank Eric Vatikiotis-Bateson for his initial suggestion for a topic that would blend work done in his lab with electroencephalography research experience from my undergraduate degree, as well as for providing me with the resources and training necessary to carry it out. Similarly, I'd like to thank Lawrence Ward for his support and advice, even as the EEG component of the thesis was abandoned due to time constraints. Finally, I'd like to thank Carla Hudson Kam for her insightful comments, as well as her professional and personal advice.

Many people beyond my committee helped with the technical aspects of data collection and analysis. Adriano Villela-Barbosa and Martin Oberg were both instrumental to setting up and manning the recording equipment during my four sessions with my singers. Adriano guided me through the first of the optical flow and correlation map analysis, in particular spending a great deal of time patiently teaching and reteaching me how to use MATLAB. Rob Furhman offered further coding guidance in the later stages of the opitcal flow and correlation map analyses. Nick Bedo walked me through the both the theoretical background of the transfer entropy analysis, and wrote a great deal of code towards its implementation.

Lastly, I'd like to thank my family and friends for keeping me (mostly sane) and healthy through this lengthy program: Evan Reznicek and Tyrone Phillips both answered some of my (myriad) MATLAB questions, proving that it really pays to keep engineers around; Andrei Anghelescu, for not only being a hilarious, supportive, and all-around awesome housemate, but also for writing me a script that I wound up not using; my cohort at UBC, for being some of the smartest, funniest, warmest, and most welcoming people I know; the various other friends I've made here in Vancouver, for turning a strange city into a place I hope to call home for a long time; Penelope Bacsfalvi, for being a wonderful mentor and encouraging presence; Natalie Weber, for being the best housemate and friend I could have wished for here (as well as for teaching me how to use LaTex); and finally, my parents. They are an inspiration and a rock, offering both emotional and financial support. I am continually grateful for the relationship we have cultivated in my adulthood, and I hope I can live at least half as generous and meaningful a life as they do.

Chapter 1

Introduction

Humans are capable of a high degree of interpersonal coordination. In some cases, coordination is largely involuntary and not necessarily a goal of the interaction, as in conversation [11, 27]. In other types of interaction, coordination is a more explicit, primary goal for all participants, such as in musical performance. Musicians attain high levels of coordination, in part through auditory feedback. Goebl and Palmer [13] examined the influence of auditory feedback on synchronization between two pianists, and found that asynchronies in the performance increased as auditory feedback decreased. Musicians may also use visual feedback, such as their co-performers' movements and gestures (shown by studies such as Clayton [9]) . Some of these movements are necessary for the production of sound (breathing, for example, or opening one's mouth to sing). Following Wanderley [32], these movements are called instrumental, or effective gestures. Motions that are not directly related to the generation of sound (such as motions made with an instrument, or head-bobbing) are called ancillary, or non-obvious gestures [32]. The current study examines the question of whether singers' instrumental and ancillary gestures and levels of coordination change over the course of the performance, and if so, how. Secondarily, it combines qualitative methods and quantitative measures to evaluate interactions between performers, specifically with regards to musical structures and gestures identified in previous studies, such as Keller [18], Williamon and Davidson [34], and Vines et al. [30].

In general in Western music, the basic form of a piece of music is determined before the musicians approach it [19, p. 207]. In this way, musical performance often requires a level of planning more similar to performing a scene from a play than spontaneous conversation. The musicians may have completely memorized a song, but if they have not rehearsed together, they still need to converge on an interpretation of the song on the spot. Convergence on a common tempo (analogous to speaking rate) and loudness is not merely incidental, but necessary. The musicians may have different ideas about the interpretation of the song, for instance, or a piece will call for changes in tempo or loudness, which must be negotiated in real time [19]. Furthermore, the musicians may make mistakes. Skilled performers are able to recover from and compensate for each other's or their own mistakes without drawing attention to them–for instance, jumping ahead in a song if their partner skips a section, altering the tempo to allow their partner a chance to catch up, or changing the pitches of their notes to remain in tune with their partner. Less skilled performers may need to stop playing or singing entirely in order to re-start the song again together, but even then, they will need to converge on a good place in the song to start over.

Like converging on a common tempo and loudness, certain musical structures appear to be more difficult to coordinate than others. Two simultaneous 'voices' or 'parts' in a musical piece may be in different keys or their time periods (meters) may not be related by simple ratios, or even related by ratios at all. These sorts of musical structures are potentially more difficult for musicians to coordinate [18, p. 24-26]. Based on the model outlined in Keller [18], singing or playing in unison would seem to be the easiest musical structure to coordinate, as the parts are rhythmically and tonally identical. Turn taking (with overlap), or aligning two contrapuntal parts (parts which have different melodies and rhythms), would be more difficult[18]. Assuming that switching between unison sections and overlapping solo sections is challenging to coordinate, musicians may struggle the first time they perform a song containing these structures.

Regardless of the challenge posed by the structure of a song, when navigating a performance in real time, performers may need to convey information about the piece of music to one another in order to coordinate their individual parts. In informal performance settings, musicians may count out loud, or verbally cue another musician to start playing: for instance, indicating to the guitarist that it is time for his or her solo. (One example of such counting behavior is Tom Petty and the Heartbreakers' performance of *Taxman* by the Beatles at the Concert for George [Harrison] in 2002 [10], where Tom Petty counts out loud at the beginning of the song to cue the rest of the band.) Thus, such communication becomes part of the performance. In more formal performance settings, this behavior can be seen as disruptive and unprofessional [17]. Furthermore, during a performance, verbal communication can be outright impossible for singers and wind instrument players, whose vocal tracts are otherwise engaged. Studies of communication between performers have shown that performers overcome this by using a wide variety of conscious and unconscious nonverbal gestures to communicate during performances. For instance, Seddon and Biasutti [26] showed a jazz musician who explicitly tapped his head to tell the rest of the ensemble to return to the main melody (or "head") of the piece.

In a study addressing audiences' interpretation of performer movements and gestures, Vines et al. [31] tested thirty musicians to establish how much information about the structure of a song was available through visual cues alone. Participants were presented either with the audio or silent video recording of a solo clarinet performance. Participants used a slider on a response panel to indicate the level of tension in the performance, as well as the phrasing of a performance. (*Tension* refers to the psychological response to a build-up of the listener's expectations of musical structure in real time, as shown in Lerdahl [21], Bigand et al. [6], Lerdahl and Krumhansl [22], and Huron [16]. *Phrasing* refers to the division of a musical piece, particularly the melody, into groups of successive notes [8].) Specifically, the participants were asked to move the slider up to indicate the beginning of a phrase and down to indicate the end. Similarly, participants moved the slider up or down to track changes in tension [31, p. 472]. Both groups were able to judge the timing of the beginning and ending of each phrase fairly accurately.

In the Vines et al. [31] study, one particular segment of the clarinet performance stood out: a long, extended note followed by a pause, where the clarinetist stopped playing for a moment before taking a deep breath and making a large swooping motion with his clarinet. Here, the phrase judgments of the two groups differed. The video-only group could see the onset of the note, but could neither see nor hear the end. The audio-only group could hear the end of the note, but did not have any visual cues to alert them to the onset. Vines et al. proposed that the video-only group judged the extended note to be held longer because they could not hear the end of the note. The audio-only group, on the other hand, lagged behind the video-only group in their judgments of the beginning of the next phrase, potentially because they did not see the clarinetist's preparatory breath and gesture [31, p. 475]. If musically-trained audiences are able to use this type of gesture to predict future actions, it stands to reason that co-performers may utilize them to coordinate their playing.

In a study examining gesture between co-performers, Williamon and Davidson [34] investigated the development of nonverbal communication between a pair of pianists rehearsing and performing duets together. The pianists videotaped rehearsals of their duets over a ten-week period, culminating in a video-taped performance. The authors examined the videos of the last two rehearsals before the performance, as well as the performance itself, in slow motion, and counted instances of eye contact between the two pianists and two types of nonverbal gestures: *hand lifts*, where the pianist would lift his hand more than usually necessary for striking the keys of the piano, and *upper body swaying*. Then, the authors identified each occurrence of a gesture or eye contact as occurring at important or less important points in each piece, based on the performers' own reported intuitions about what parts of the piece were important, in order to ascertain whether gestures and eye contact occurred more often at more important parts of the pieces.

Analysis of variance revealed that the number of gestures at the key points that the pianists identified increased in the videos approaching the performance. The authors claim that the level of synchronization between the performer's gestures also increased over the course of the rehearsals and performances, but no quantitative analysis was performed to support this claim [34, p. 60-61]. Williamon and Davidson also found that eye contact between the two pianists increased significantly over the course of the last two of the four rehearsals and the performance [34, p. 61-62].

Level of musical experiences and familiarity between co-performers also affect the form of nonverbal communication during performance. Ginsborg and King [12] examined the gestures and eye contact of four singer-pianist duos who had been playing together for some time. Two pairs were professional, and two pairs were students. From a subset of these musicians, they formed two pairs of singers and pianists who were new to each other and had different levels of expertise. Ginsborg and King recorded the rehearsals for each pair, and then categorized the gestures and movements made as *states*, or actions with durations (such as gazing between performers); *points*, or actions without durations (such as a glance from one performer to another); and *gestures*, or all movements that didn't fall into another category, such as *pulsing* (bobbing head, hand, or body to the beat), *shaping*, *conducting*, gazing, or glancing. They found that when performers rehearsed with familiar partners, or with partner at the same level of expertise, they used a greater variety of physical gestures, and to a greater extent, than performers who were rehearsing with a new partner of a different level of expertise [12].

Although the aforementioned studies do offer quantitative analyses of amount and duration of performers' movements, they do not closely examine the spatiotemporal relationships between them. However, a few other studies outside of music performance have looked at methods of quantifying these spatiotemporal relationships between motion signals. Video is comprised of pixels arranged in columns and rows. Optical flow calculates the amount and direction of pixel movement between frames in greyscale video [15]. Optical flow analysis does not require trackable markers. Therefore, it is more non-invasive than motion capture analysis while still providing the same level of spatiotemporal information [3]. The algorithm put forth in Barbosa et al. [4] calculates displacement vectors for each pixel in video. The array of displacement vectors for all pixels is called the optical flow field [4]. To find the total amount of motion in each frame step, we can calculate the sum of each vectors magnitude. Furthermore, the user can home in on just the gestures they want by designating *regions of interest* (ROIs) within the larger camera frame. [3]

The signals optical flow analysis yields can then be analyzed in relation to one another using methods such as correlation map analysis (CMA), which calculates the *instantaneous correlation coefficient* between two time series within a range of temporal offsets, and renders it into a two-dimensional *correlation map* [4]. Correlation map analysis has been used in linguistic studies on both individual and pairs of speakers, such as in Vatikiotis-Bateson et al. [29] and Barbosa et al. [4], which lays out an algorithm that calculates the instantaneous correlation coefficient between two signals (visual, audio, or both), and used it to look at the coordination between facial movement, hand gestures, and speech acoustics in Plains Cree. Results showed strong correlations between the root mean square (rms) amplitude of the speech and the movements of the face and hands of just one speaker. Real synchronization is relatively rare in coordinated behavior, regardless of how musicians may strive for it [3, p. 1], so the algorithm allows the user to set a range of temporal offsets over which the algorithm will calculate correlation. The algorithm also allows the user to set the size of the filter window, or in other words, how quickly past samples are forgotten [3, p. 4]. Different sized filters will be more or less sensitive to faster or slower changes in correlation between the two signals. The resulting 2D correlation map plots the degree and direction of correlation for each temporal offset over time. The filter is bidirectional–that is, it takes into account both past and future samples.

Transfer entropy analysis, like correlation map analysis, provides another method of examining the relationship between two time series [25]. Whereas CMA (as laid out in Barbosa et al. [4]) works bidirectionally, taking into account both past and future samples, transfer entropy analysis examines how predictable the values in one time series are based on the previous values both in that same time series and in the other time series being analyzed, and whether there is a directional, or causal, relationship between two signals. In this context, causal means that the future of one signal is more easily predictable based on past samples from the other signal as well as its own [33]. It is not meant to imply with any level of confidence that one singer is leading the other. Because it allows for non-linear causal interactions–as in studies examining the amount and direction of information being transferred between brain regions [5]–it may be well-suited to examining relationships between singers gestures during performance.

The present study combines the qualitative methods presented in the music performance literature with the quantitative methods outlined above to investigate whether level of correlation corresponds to the frequency and duration of gestures such as eye contact and head movement between musicians during rehearsal and performance. Do greater durations and frequencies of gestures correspond to higher levels of correlation between singers, or may one be present without the other? Relatedly, does the level of correlation between singers increase over time the way amount and duration of gestures seems to do? Do different types of gesture impact levels of correlation differently?

In the present study, I followed Williamon and Davidson [34] and Ginsborg and King [12] in recording video and audio for pairs of singers performing duets together for the first time; that is, with no prior rehearsal. I then annotated the video and audio to identify different types of movements and gestures. This included phrases sung together or alone (*Unison* and *Solo*), breaths before unison and solo sections (*Unison Breaths* and *Solo Breaths*); glances directly at the other singer (*Looks At*); glances up, so they could plausibly see the other singer in their peripheral vision (*Looks Up*); and times when singers made mistakes. (A table outlining these types of gestures is shown in Table 2.3.)

Using these annotations, I examined the total duration of each type of token to determine how they changed over the course of the recording session. I then used optical flow analysis to calculate motion vectors from the video data for each singers head and body. Then, I examined these signals using correlation map analysis and transfer entropy analysis to determine the extent to which the interaction between the regions of interests for each individual singer and between the two singers changed over the course of the recording session. Correlation map analysis was used to calculate the levels of correlation between pairs of motion signals, as well as the root-mean-squared amplitude of the audio, both within singers and between singers. The correlation results were used to determine if the levels of correlation change over the course of each session, both for entire songs and for solo and unison segments. They were also used to examine if there was any difference in correlation levels between the solo and unison segments. Transfer entropy analysis was used on the motion signals to identify if there is information transfer between the two signals during time windows for the different types of movements and gestures examined, and if so, at what temporal lag.

Following Williamon and Davidson [34], I expected to see either increased number or duration of looks up by both singers over the course of each recording session. Similarly, I hypothesized that if levels of correlation are related to familiarity between singers (as increased levels of gesture between musicians seem to be, according to Ginsborg and King [12], then the later duets in each recording session would show higher and more predictable peaks in correlation than the earlier duets. Finally, I predicted that if musical events such as breathing and the onset of unison versus solo singing, and gestures such as eye contact, convey important information between performers, then we would see higher levels of correlation and transfer entropy around these moments in the performance.

Chapter 2

Methods

2.1 Recording Methods

2.1.1 Subjects

Five singers were recruited to participate in this study. Each singer filled out a short survey (Appendix A) detailing their prior musical training and experience, as well as their post-recording impressions of their performance and any non-verbal communication techniques they were aware of employing.

- Singer 1 (M, 23) performed in Sessions 1, 3, and 4. He identified voice as his primary instrument, reporting that he had 10 years of experience singing and eight years of experience performing in ensembles. In addition, he reported that he had received three years of piano lessons and also had six years of lessons and ensemble performance experience with French horn.

- Singer 2 (F, 22) performed in Session 1. She reported that voice was her primary instrument, having 10 years experience singing, including seven years of ensemble performance. She also had 12 years of experience with piano (11 years of lessons).

- Singer 3 (F, 25) sang in Sessions 2 and 3. She reported voice as her primary instrument, with one year of individual lessons and two years of choir. She also reported taking piano lessons for ten years, and playing drums with an ensemble for three years.

- Singer 4 (M, 26) performed in Session 2. He had no formal training in any instrument, including voice, and no experience performing in an ensemble.

- Singer 5 (F, 24) performed in Session 4. She reported that voice was not her primary instrument. However, she had played alto saxophone with an ensemble for between five and seven years. She had not sung with a choir in over 10 years, and had not played music with a group of any kind for several years. For this reason, she was deemed a less experienced singer than other singers in this study with similar amounts of ensemble performance experience.

These five singers were paired into four pairs, and recorded in four sessions, as shown in Table 2.1:

Table 2.1: Sessions and Singers

Session	Singer A	Singer B
Session 1	Singer 1	Singer 2
Session 2	Singer 3	Singer 4
Session 3	Singer 1	Singer 3
Session 4	Singer 1	Singer 5

Each session was analyzed apart from the other sessions. Singer 1 and Singer 3 participated in multiple sessions, for ease of subject recruitment. This created a potential imbalance in familiarity with the songs between singers within the later singer pairs, as well as with the structure and demands of the study itself. For this reason, each session should not be viewed as fulfilling the same set of criteria (two singers who have never performed the songs before, either together or with another partner).

2.1.2 Duets

In each session, the singers sang the same three duets (**??**): (*California Dreamin'* by the Mamas and the Papas (Phillips and Phillips [24]), *The Battle of Evermore* by Led Zeppelin (Page and Plant [23]), and *Scarborough Fair/Canticle* by Simon and Garfunkel (Traditional and Simon [28])) twice in a-b-c-a-b-c order (see Table 2.2). The three duets were chosen because they are not romantic, and because they contain phrases sung both in unison and alone. Each song was written for two parts. One part comprised the main melody and the bulk of lyrics, and could therefore conceivably be referred to as the lead part. Each part occupied approximately the same tonal range, and could reasonably be sung by either singer in a male-female pairing. The singers were allowed to decide between themselves who would sing what part. However, in all pairings, the male singer sang the lead part.

Table 2.2: Order of Duets and Abbreviations

Order	Duet Title	Abbreviation
1	California Dreamin'	CD1
2	Battle of Evermore	BE1
3	Scarborough Fair	SF1
4	California Dreamin'	CD2
5	Battle of Evermore	BE2
6	Scarborough Fair	SF2

Of the three, *California Dreamin'* was the simplest, because it contained lyrics repeated between the two parts, and was potentially the most familiar of the three songs. *Scarborough Fair/Canticle* was also likely to be easier, because half of the lyrics and the melody are from a traditional English song (*Scarborough Fair*), and it was performed by a well-known folk duo.

However, the two parts were the most independent from one another, compared to the other two duets. *Battle of Evermore* was the most challenging, because it was the least familiar to the singers, and had the most unrepeated lyrics.

Before coming into the lab, the singers were provided with recordings of the three songs, as well as sheet music for each duet, which I transcribed from the recordings. The singers chose to use sheet music, except for the first four duets sung in Session 2 (the only pair that did not include Singer 1).

The singers settled on a comfortable key for each duet immediately prior to singing that duet on the day of recording. In order to avoid synchronization to an external beat, they did not use a metronome to select a tempo. Although they were not given explicit instructions regarding ”counting [themselves] in”, none elected to verbally count out the beat at the beginning of the duets.

2.1.3 Recording

Video was recorded using a Panasonic AG-DVX100BP camera at 29.97 frames per second. Audio was recorded at 32 kHz, with separate channels for each singer, using lavalier microphones mounted on a head-rig approximately 20 cm above and to the left of the mouth. Video was converted from DV (48kHz, 29.97 frames per second) to H264 format (32 kHz, 29.97 fps) in Quicktime 7. Each duet was recorded separately from the others, and then further trimmed in Quicktime 7 to include only a few frames immediately before the onset and end of the song. The individual audio channels (one for each singer) were exported from the video as separate wav files (also in Quicktime 7).

2.2 Annotation of Audiovisual Recordings

Williamon and Davidson [34] and Wanderley [32] outline a few different types of motions and gestures that musicians might use to signal the structure of a musical performance. These might be intentional gestures such as direct eye contact and specific hand gestures (see Williamon and Davidson [34]) or incidental to the act of making sound, like taking a breath before the beginning of a musical phrase (see Wanderley [32]). Following these two studies, I annotated: sections where the singers sang in unison (Unison), sections where the singers sang alone (Solo), sections where the singers breathed before both unison and solo sections (BreathU and BreathS, respectively), sections where the singers looked directly at each other (LooksAt), and sections where the singers could reasonably see one another in their peripheral vision (LooksUp). Following Williamon and Davidson, I expected to see either an increased number or duration of looks up between singers over the course of the recording session.

Using ELAN [1], I annotated the video to identify times when the singers looked up or at one another. ELAN is a software tool developed at the Max Planck Institute for annotating video. It allows the user to create tiers (either hierarchically linked, or independent) that contain annotations of events within a video like gestures or transcriptions of speech [20]. In the present study, I created an independent tier in ELAN for for each singer. Each tier contained both LooksAt and LooksUp (see Figure 2.1). I identified these gestures without the audio, in order to avoid auditory bias in annotation, in case hearing the singers might influence judgments on

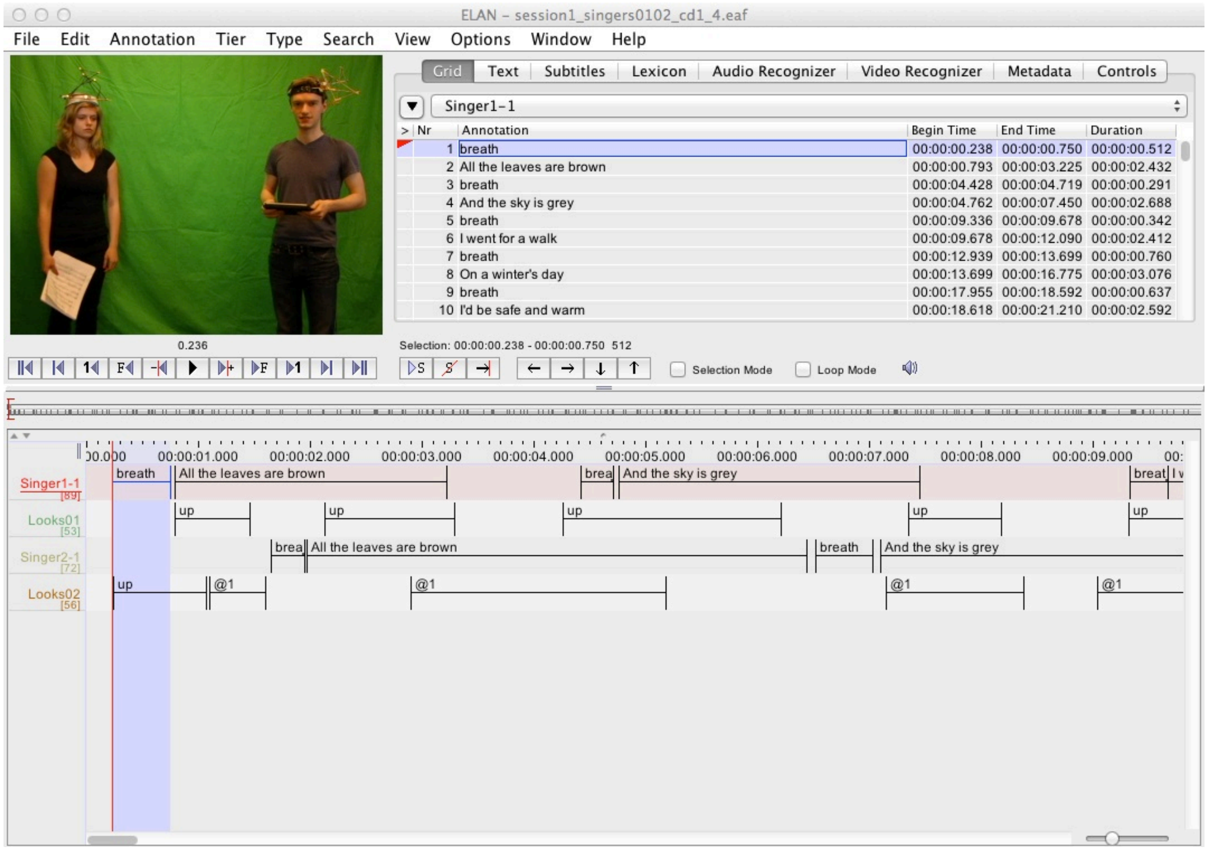


Figure 2.1: An example of annotation using ELAN, with video in the upper left corner, the text of the annotations listed in the upper right, and the lower half showing the four tiers containing annotations. The interface allows the user to navigate the video frame by frame, aiding in precise annotation. The side-by-side tier view allows the user to see where gestures such as glances overlap in time with the text of the song

when singers moved. ELAN allows the user to examine video frame-by-frame. Annotations for looks began at the frame where the singer either opened their eyes or began turning to look either up or directly at the other singer, and ended when the singer closed their eyes or began turning down or away.

I used Praat [7] to annotate the audio, and to identify times when the singers began singing in unison or solo, and when they breathed (where breath was both audible and visible in the waveform). Praat is a software that allows the user to examine an audio waveform at high resolution, and transcribe that audio into text grids, which are exportable as files that can be used with programs such as ELAN. I recorded these annotations in two independent text grids, and imported the text grids to ELAN. Finally, I exported all four tiers from ELAN to a text file which included the singer and the type, beginning, end, and duration of each gesture. I then sorted these time windows by type (Unison, Solo, BreathU, BreathS, LooksAt, and LooksUp) and singer.

Lastly, I annotated moments where the singers made mistakes. I defined mistakes as moments where one or both singers sang the wrong lyrics and then repeated the line again with the correct ones, or paused where a pause was not written in the score.

Table 2.3: The types of gestures annotated using Praat and ELAN, with their descriptions

Gesture	Description
Unison	Singers sing the same thing at the same time
Solo	One singer sings a distinct phrase by themselves
Unison Breaths	Breaths taken before the onset of a phrase sung in unison
Solo Breaths	Breaths taken before the onset of a phrase sung alone
Looks At	One singer looks directly at the other singer
Looks Up	One singer looks up, with the other singer visible in their peripheral vision
Mistakes	One or both singers sang the wrong lyrics and then repeated the line again with correct lyrics, or paused where there was not a pause written into the score

2.3 Optical Flow Analysis

In the present study, I designated each singer's head and body as ROIs of about the same width. In the case of the head ROI for each singer, I measured from the top of the shoulders to the top of the head. For the body ROIs, I measured from the shoulders to roughly the top of the hips. All ROIs captured the full range of horizontal and vertical motion of the head or body over the course of the video. Figure 2.2 shows the user interface as ROIs are selected for a particular duet in Session 1. The user may view the entire clip to confirm that the selected ROIs contain the full range of motion for the body part they intend to capture–in this case the singers' heads and bodies.

The algorithm computes five different measures for each of the four regions of interest: total magnitude of movement (mag), horizontal movement (x), magnitude of horizontal movement (xmag), vertical movement (y), and magnitude of vertical movement (ymag). The These signals are time series that reflect changes in motion over time. I performed two different types of analysis on these signals, as well as the rms amplitude of the audio: correlation map analysis (for all signal pairs) and transfer entropy analysis (for motion signal pairs only). Figure 2.3 shows an example of the summed magnitudes for both singers' heads in a particular session and duet, plotted over time in frames on the x axis. Time series like these can then be analyzed using techniques such as correlation map analysis and transfer entropy analysis.

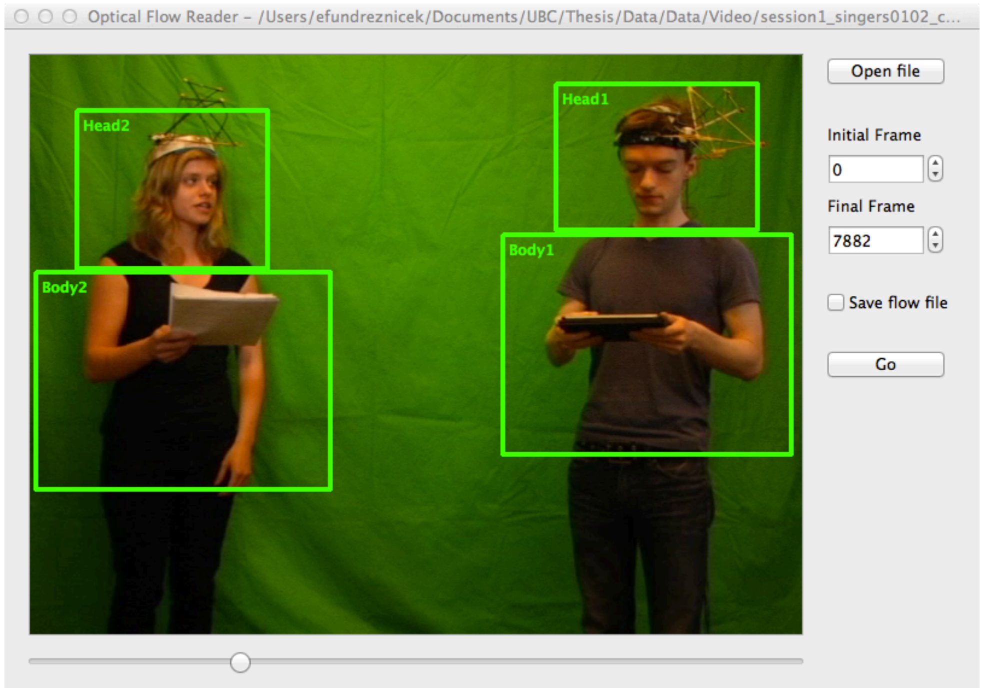


Figure 2.2: The above screen capture shows the designated regions of interest (ROIs) for a particular duet in Session 1. A separate ROI is specified for each singer's head and body. The user is able to view the entire video file to confirm that each ROI captures the entire range of motion of the singer's head and body.

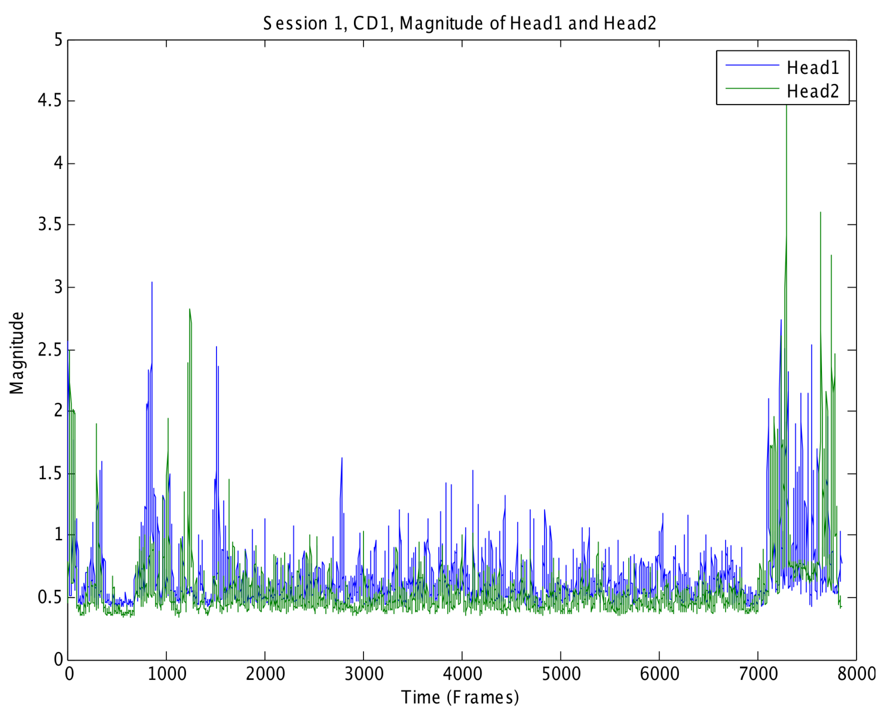


Figure 2.3: An example of the series of summed magnitudes for both singers' heads in Session 1, California Dreamin' 1, over time in frames on the x axis.

2.4 Correlation Map Analysis

For each session and duet, I first normalized all signals (motion and audio) to the same total number of samples. Using the MATLAB Toolbox for Audiovisual Speech Processing [2], I then used the algorithm to calculate the instantaneous correlation between 15 different pairs of signals (see Table 2.4), using two different values of η (η=0.310 and η=0.031, as per Barbosa et al. [4]), for each measure (mag, x, xmag, y, and ymag) and for two different temporal offset ranges (-1 to 1 seconds and -0.5 to 0.5 seconds, or Δ=1 and Δ=0.5), following Barbosa et al. [4].

Table 2.4: Within-Session Signal Pairs

Within Singer	Between Singers		
Head–Body	BodyA–BodyB	HeadA–BodyB	rmsA–BodyB
rms–Body	BodyA–HeadB	HeadA–HeadB	rmsA–HeadB
rms–Head	BodyA–rmsB	HeadA–rmsB	rmsA–rmsB

Figure 2.4 shows an example of a correlation map. The top three graphs show the amplitude of the audio and the summed magnitude of both the signals being analyzed for correlation. The bottom graph is a heat map which shows the level of positive and negative correlation (from -1 to 1, shown in the color bar to the right) across temporal offsets from -1 second to 1 second (along the vertical axis), over time measured in frames (along the horizontal axis).

For each pair, measure, η-value, and Δ value, in addition to plotting a 2D correlation map, I gathered the total number of samples ("hits") where the correlation value *p(k)* was above 0.5 (positive correlation) and below -0.5 (negative correlation) at each temporal offset, where *p(k)* accounts for 25 percent of the variance (following Vatikiotis-Bateson et al. [29]).

I then gathered the histogram values of each duet for each session, signal pair, and value of η and Δ. I divided each value in each duet by the total number of samples in that duet, in order to account for the difference in duration between duets. This returned the percent number of correlation hits for the entire duet. Then, I plotted the number of correlation hits for each duet over each temporal offset. Comparing the levels of correlation for each duet across the entire session reveals changes in the overall levels of correlation, as well as differences in the levels of correlation at different offsets between duets. This is one way to determine if the overall level of correlation between singers increased over the course of the recording session.

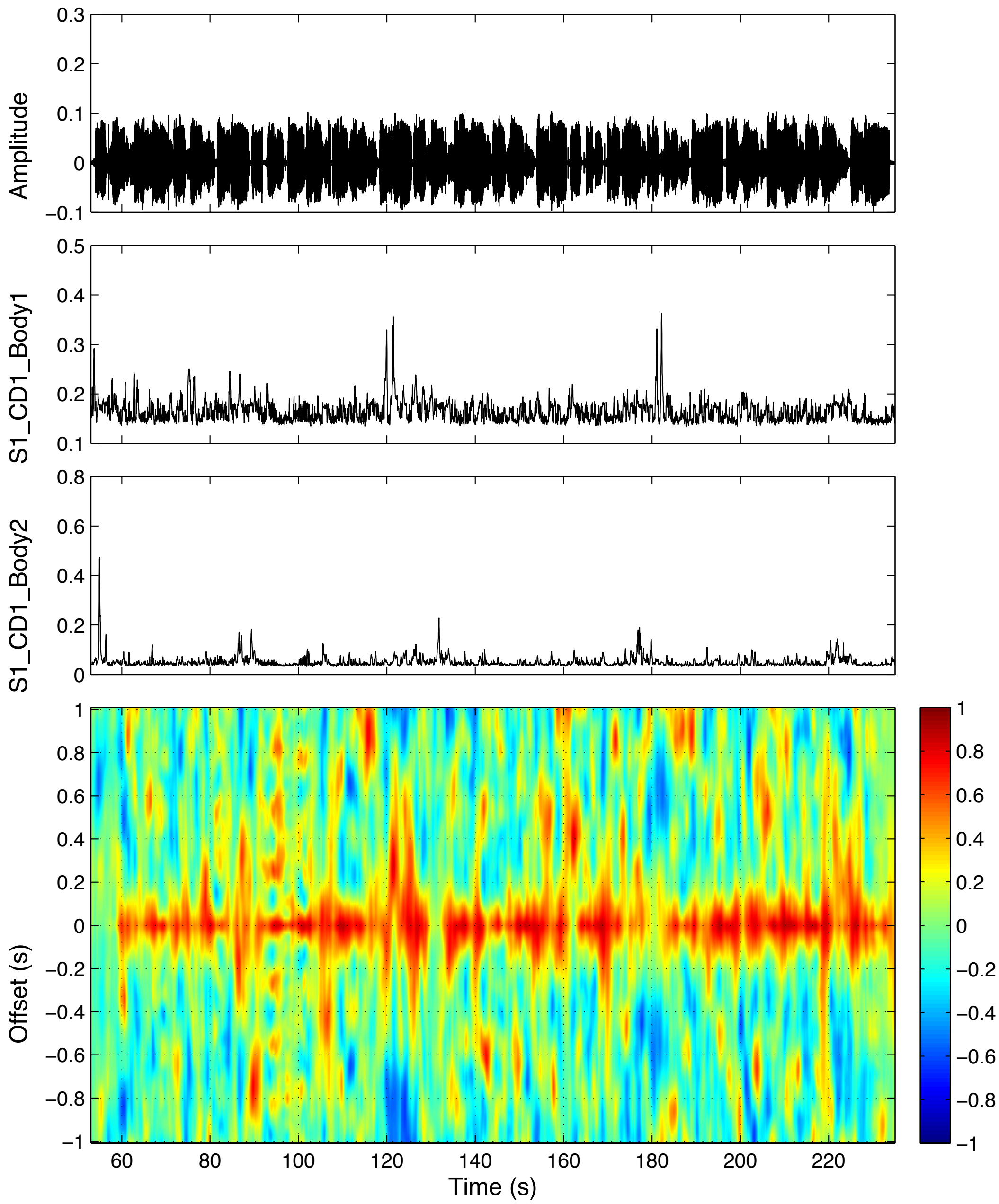


Figure 2.4: An example of a correlation map for Session 1, California Dreamin' 1. The top graph shows the amplitude of the audio signal. The second graph from the top shows the magnitude of Singer 1's body motion. The third graph from the top shows the magnitude of Singer 2's body motion. The bottom graph shows the continuous correlation coefficient over time (along the horizontal axis) across a range of temporal offsets from -1 to 1 second along the vertical axis.

2.5 Transfer Entropy Analysis

In the present study, I drew on the annotations I made in the video and audio to divide each duet in each session into six types of tokens: Unison, Solo, BreathU, BreathS, LooksAt, and LooksUp (see Table 2.3). I included tokens for both singers in each session. So, for example, Session 1 breaks down into six token types for both Singer 1 and Singer 2.

For each token, I drew a window from 150 frames before the onset of the token, to 900 frames after the onset–longer than the length of the longest token. I then pulled the data for these time windows from the original motion signals. Using the TIM MATLAB toolbox (Gomez-Herrero et al. [14], found at http://www.cs.tut.fi/∼timhome/tim/tim.htm), I computed the transfer entropy on the raw motion signals for each token type in each duet for the magnitude measure and the six motion signal pairs (from the 15 pairs used in the correlation map analysis, excluding pairs containing rms amplitude).

Each signal pair yielded two sets of results-one set for each direction (from signal 1 to signal 2, and from signal 2 to signal 1). Similar to the temporal offsets used in the correlation map analysis, transfer entropy was calculated at lags up to 0.25 seconds (or roughly 15 frames). Note that, whereas the correlation map analysis covered a wider range of temporal offsets (up to $\pm$ 15 seconds), and therefore the samples in one signal before and after the samples in the other, the transfer entropy analysis only takes into account the samples before. The visualization of the transfer entropy results is somewhat similar to the 2D correlation maps, in that it plots the amount of information transferred from one signal to the other at each lag (from 1 to 15 frames prior) over time (the -150 to 900-frame time window mentioned above). It differs in that it calculates the information transferred in a particular direction from one signal to the other (whereas the correlation map analysis performed in this study was explicitly non-causal and makes no claims about directionality).

Chapter 3

Data Analysis and Results

3.1 Results of Annotation of Audiovisual Recordings

The annotations for times when the singers sang in unison or solo, and when they breathed before singing, were used to constrain correlation map and optical flow analyses. The rest of the tokens fell into three categories: LooksUp (where one of the singers looked up and could plausibly see the other in their peripheral vision), LooksAt (where one singer was looking directly at the other), and Mistakes (where one or both singers sang the incorrect words, hesitated to sing the next words, or started a line over). I calculated both the total number of tokens of each type and the total duration of each token type as a percentage of total song duration. Figure 3.1 and Figure 3.2 show comparisons of the number of tokens and the percent duration over the course of each recording session. The top chart shows the number of tokens for each singer in each recording session (for instance, S3-Singer 1 indicates the number of tokens of, for instance, Looks Up, by Singer 1 in Session 3). Number of tokens is plotted along the y-axis, and the duets are plotted in order of occurrence along the x-axis.

The number and duration of looks at and up did not increase over the course of any sessions. Certain singers seemed more prone to looking up or at one another than other singers (for example, Singer 3 and 4 in Session 2. Singers made the most mistakes during BE1, possibly because it was the first run-through of the longest, wordiest, and least familiar song. Overall, the singers did not look up or at each other more or longer during the later duets. They did look up and at each other less during harder duets, possibly because they were focusing on their sheet music. This potentially counters Williamon and Davidson's [2002] findings, although the recordings in this study are far shorter, and presumably the singers in the present study achieved less proficiency by the end of their single recording session than Williamon and Davidson's [2002] pianists did at the end of theirs.

Each pair of singers made the most mistakes during the first round of *Battle of Evermore* (BE1), which was presumably the least familiar duet of the three, and possibly the most complicated, at least in terms of lyrics. In Session 1, the singers primarily made mistakes in the first three duets. The singers in Session 3 and 4 made no mistakes, possibly because in Session 3, both singers had participated in previous sessions, and in Session 4, one singer had participated in two previous sessions. Session 2 stands out in terms of mistakes, possibly because the singers elected to try to sing without sheet music for the first four duets. Because of this,

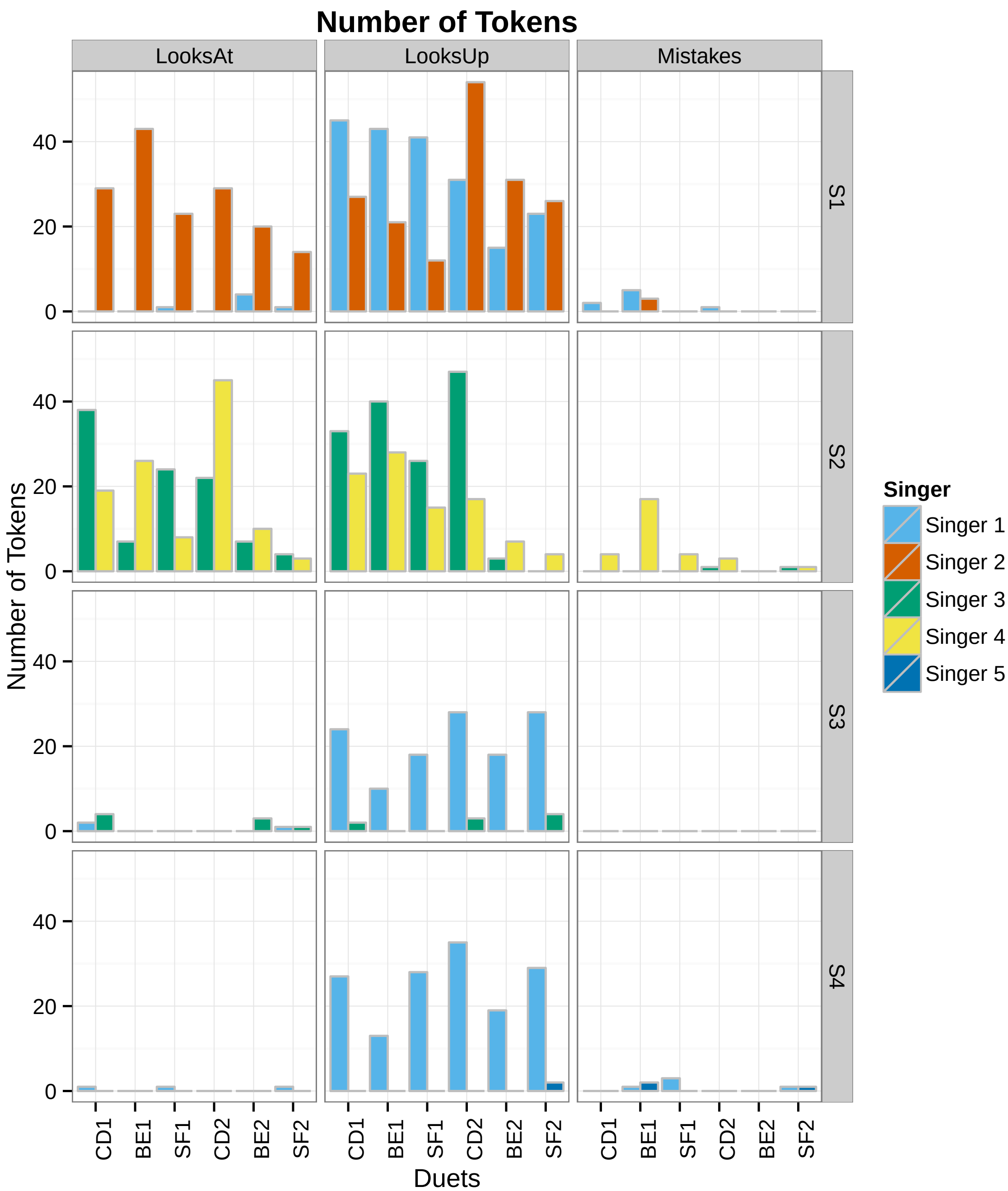


Figure 3.1: Number of Tokens

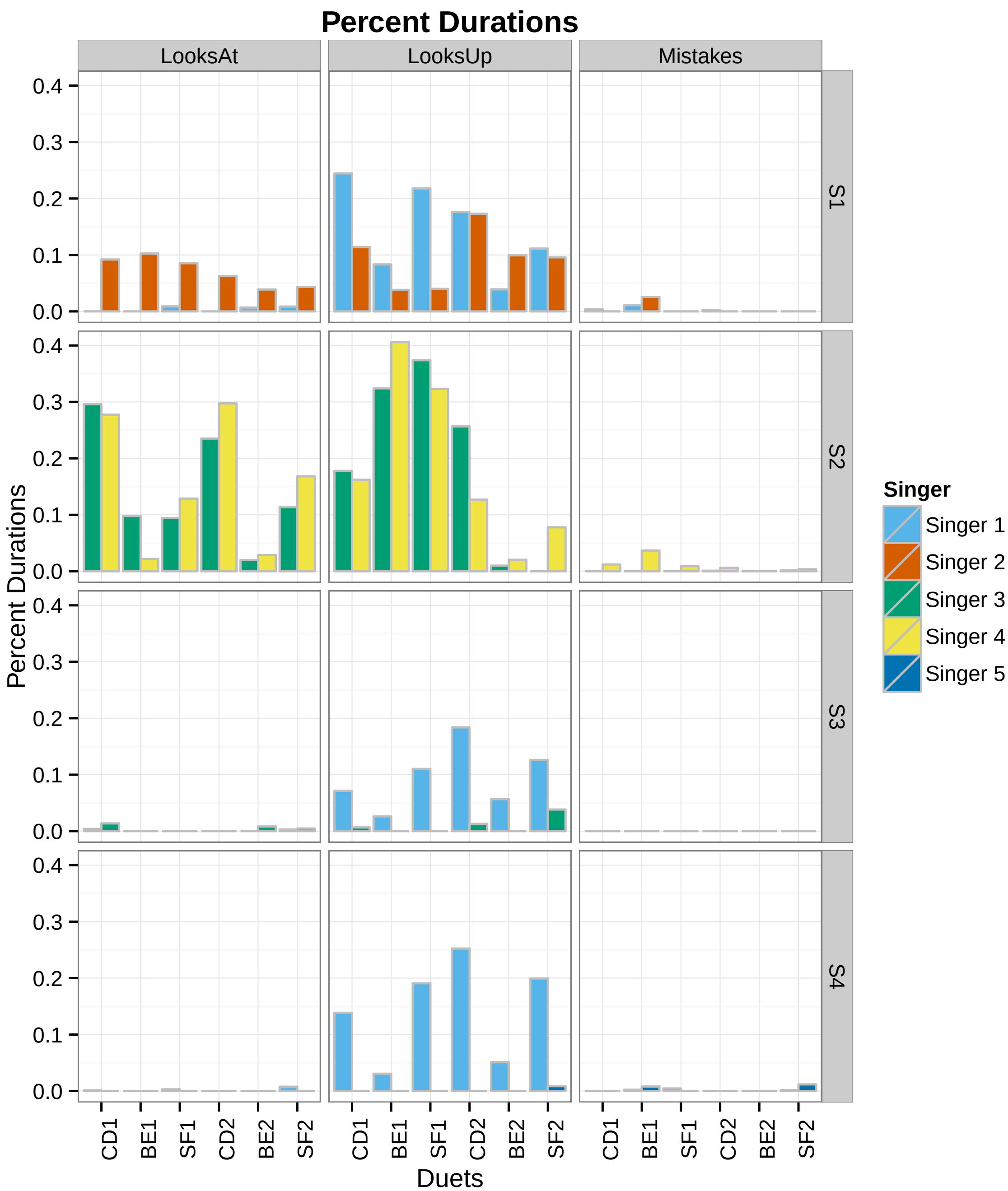


Figure 3.2: Percent Durations

Singers 3 and 4 were at a distinct disadvantage compared to the other three pairs. They made by far the most mistakes in BE1, but still made a few mistakes (approximately 5 per duet) in the last two duets.

In Session 2, Singer 4 made the bulk of the mistakes. This disparity between Singer 3 and Singer 4 is possibly due to the difference in their roles in the duets; in all three duets, Singer 4 sings what could conceivably be called the lead part. He is also a less experienced singer than Singer 3. In songs where the singers did not use sheet music, he often elected to improvise lyrics rather than attempt to recall the original ones, possibly to avoid long pauses. Because Singer 3's part often consisted of repeating lyrics that Singer 4 had just sung previously, Singer 3 chose to repeat the lyrics Singer 4 improvised, rather than pause to recall the written lyrics herself. These did not count as mistakes in the annotation, even though Singer 4 was not singing the written lyrics.

3.2 Correlation Map Analysis Results

3.2.1 Whole Song Correlations

I focused on the summed magnitude (mag) for Sessions 1 and 3, as the singers in those two sessions were most evenly matched in terms of musical training and the rehearsal of these specific duets. The singers in Session 1 had never performed the duets before, and the singers in Session 3 had each participated in one previous session, although not with each other.

The signal pairs can be categorized three ways–within-singer motion signal pairs, between-singer motion signal pairs, and rms-amplitude pairs. For both within- and between-singer motion signal pairs, negative correlations generally peaked before -0.10 seconds offset and after .10 seconds offset. For all signal pairs, positive correlations peaked at around zero offset, and peaks were more loosely grouped when η=0.0310747 than when η=0.31003. The larger η-value corresponds to a shorter sampling window, and a higher sensitivity to local changes in correspondence between the two signals. [3] Also for all signal pairs, later duets did not consistently show higher levels of correlation, even when normalized for duet duration.

Figure 3.3 is an example of between-singer motion signal pairs. Specifically, it shows the results for Session 1, Body1-Body2 and Session 3, Body1-Body3, where *Delta* refers to the size of the offsets on either side of 0 (from -1 to 1 second). For Session 1, the four graphs show the percent number of hits (on the y-axis) across the range of offsets (on the x-axis) for each duet in the session. The top two graphs show the percent of negative correlation hits where η=0.0310747 (on the left) and η=0.31003 (on the right). The bottom two graphs show positive correlation hits for the same two η values. For both negative and positive correlations, for both values of η, levels of correlation did not increase in later duets as compared with earlier duets.

Correlations between the two rms amplitude signals patterned much the same as the motion signal pairs, with negative correlations peaking before -0.10 seconds and after 0.10 seconds. Positive correlations peaked around 0 offset. (See Figure 3.4.) Correlation peaked around 0 seconds offset, between 60% and 80% number hits.

For within-singer rms amplitude-motion signal pairs, there were some small peaks. In Session 3, rms3$\sim$Body3 showed peaks in negative correlations around 0 for SF1 and SF2, and for BE2. Also in Session 3, there were peaks for the rms1$\sim$Head1 pair in negative correlations

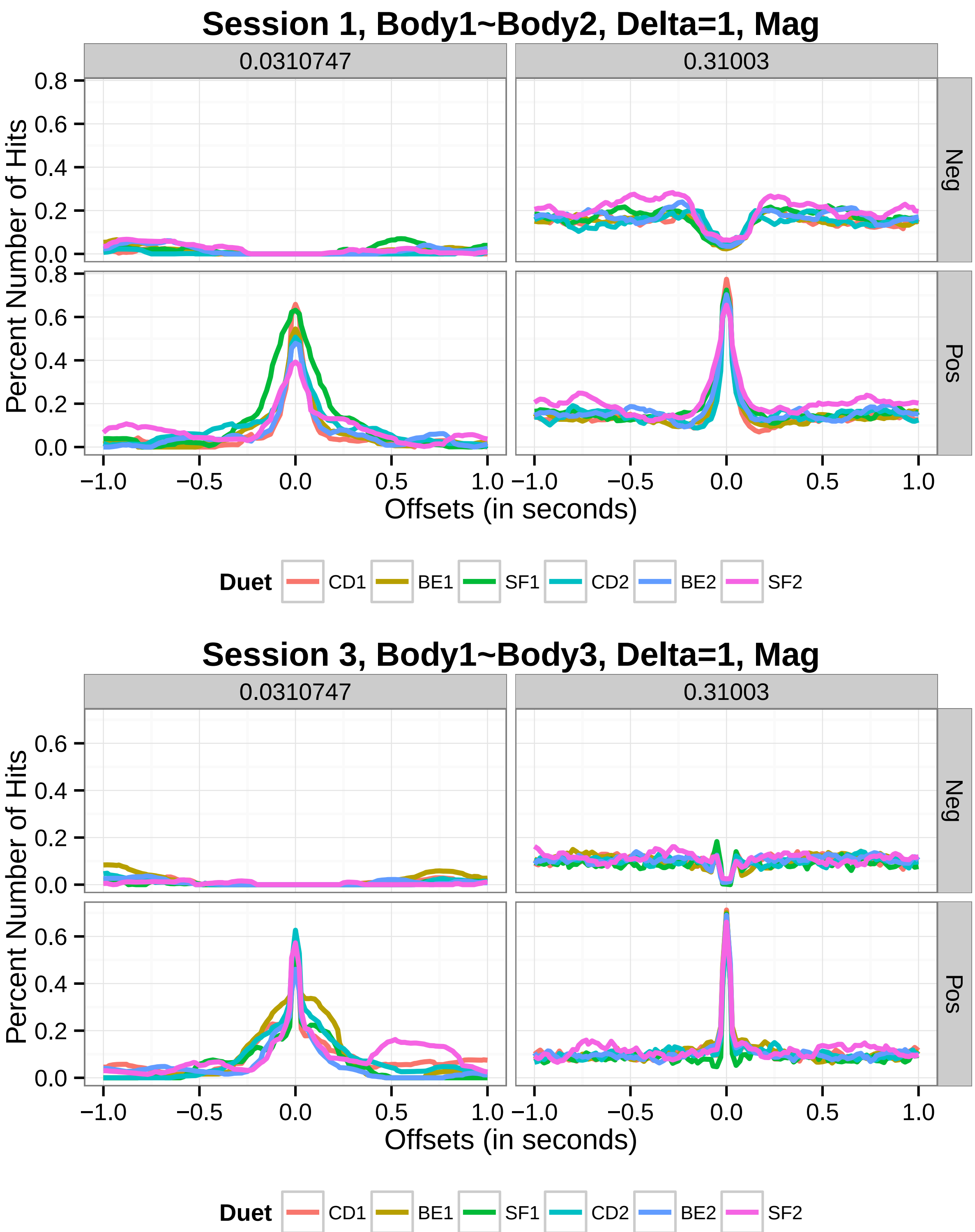
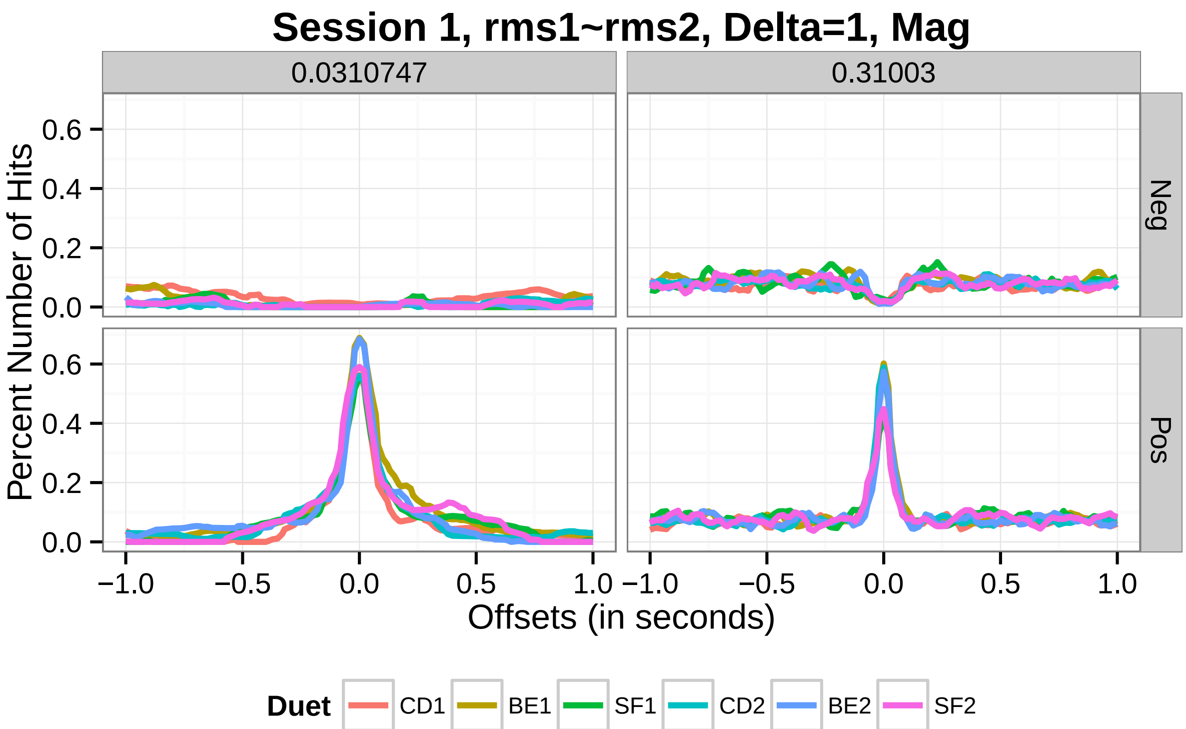


Figure 3.3: Between-singer body motion signal pairs for Session 1 and Session 3. Each graph shows the percent number of hits plotted across temporal offsets (from -1 to 1 seconds), for both positive and negative correlation hits, and for both values of η. *Delta* refers to the offsets on either side of 0.



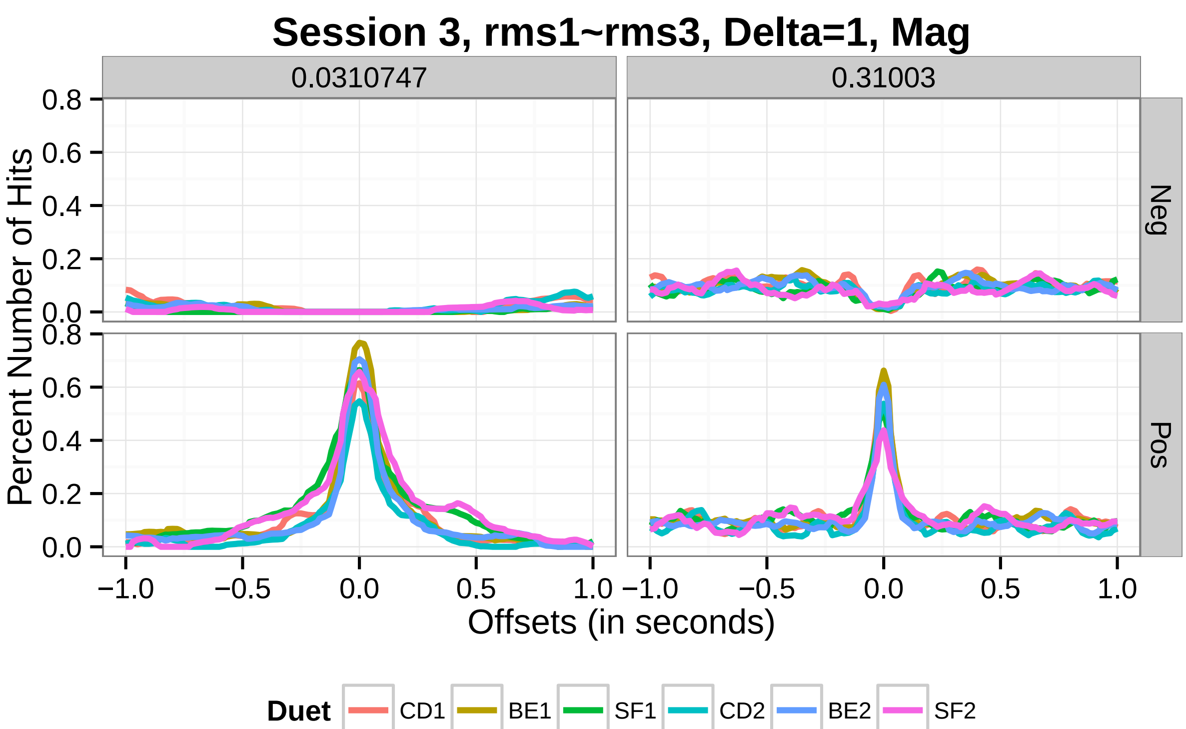


Figure 3.4: Between-singer rms amplitude signal pairs for Session 1 and Session 3.

around 0 offset for BE1, SF1, and SF2. rms1~Body1 showed a similar pattern. In general, rms~Head pairs showed a higher percentage of hits than rms~Body pairs.

For cross-singer rms amplitude pairs, there was some peakiness, but the percent number of hits never exceeded 20 percent. In contrast, within-singer motion signal pairs peaked around 60 and 70 percent number of hits, and across-singer motion signal-pairs peaked more between 10 and 50 percent number of hits.

3.2.2 Unison Correlations

The low correlation levels in within-singer rms-motion pairs might have been due to the inclusion of solo sections in the whole-song correlation maps. To rectify this, I isolated the significant positive and negative correlation hits for each unison segment in each duet. I then summed the correlation hits for the unison segments, and divided the sums for each duet by the total number of samples in the unison segments of that duet. I compared these in the same manner as I compared the percent of correlation hits for the entire duet segments.

First, there was no pattern suggesting correlation levels increased over the course of the recording session. There were high levels of within-singer positive correlation around zero offset. There was also high BodyA~BodyB correlation (see Figure B.7), but not generally in the later duets. Importantly, there were still similar patterns for rms-amplitude/motion signal pairs. Overall, they showed lower levels of correlation, and fewer peaks.

A few signal pairs did show higher levels of correlation for the last three duets. In Session 3, the rms1~rms3 pair showed high positive correlations around zero offset for CD2, BE2, and SF2 when η= 0.0310747, but not when η=0.31003 (see Figure B.11). Body3~Head3 (**??** and Head1~Head3 (Figure B.8)showed high levels of correlation for BE2 and SF2. Body1~Body3 (Figure B.7) showed highest peaks for SF2.

3.2.3 RMS Amplitude Pairs and Larger Offset Windows

The lack of distinct correlation patterns for rms amplitude signal pairs, especially within singer, contradicts previous findings from Bateson et al. (cite) This may have been because initially I looked at small offset windows–between -0.5 to 0.5 seconds, and -1 to 1 second. Because the size of each musical phrase is much longer than one second, I also calculated correlation values for -5 to 5 second and -15 to 15 second offsets.

In general, for both the plus/minus 5 second offsets and the plus/minus 15 second offsets, peaks did not exceed 15 percent number of hits. However, they did occur at more regular intervals outside of the -1 to 1 second offset windows. (See **??**.) There was still no definitive pattern showing greater levels of within-singer correlation in later duets; instead, it seemed primarily tied to the particular duet being sung. For Session 1 at least, the rms1 and rms2 amplitude pair continued to show the highest percent number of correlation hits near to 0 seconds offsets-between 50 and 60% hits.

3.3 Transfer Entropy Analysis Results

Transfer entropy analysis relies on large sample sizes (n 30) to generate significant results. In the present study, in general, each token type consists of less than 30 tokens. Because of limitations of sample size on further analysis, I have chosen to give a general overview of the results as represented in heat maps. Figure 3.5 shows an example of a heat map for Solo Breaths in Session 1, Battle of Evermore 1, from Body1 to Head1, calculated over tokens derived from annotation of Singer 1. The heat map shows the level of transfer entropy (shown in the color bar on the right) at different frame lags (plotted vertically) over time in frames (plotted horizontally). This corresponds roughly to the negative offsets in the correlation maps (for instance, from 0 to -15 frames, or roughly -0.5 seconds).

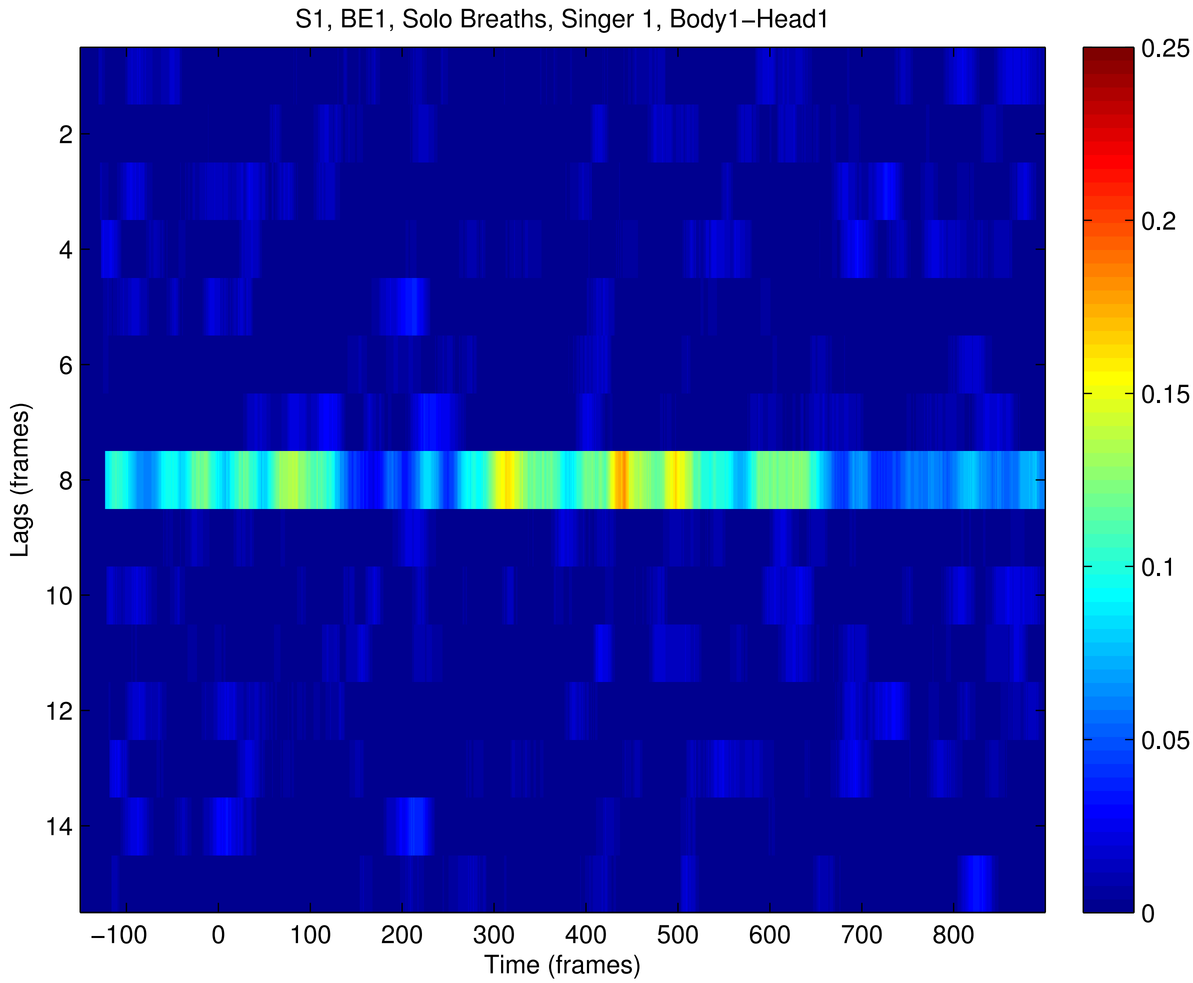
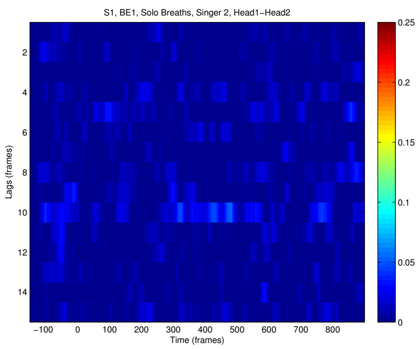
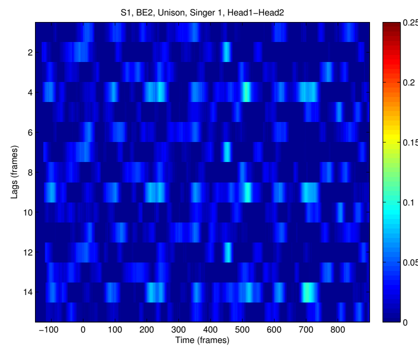


Figure 3.5: Example Heat Map

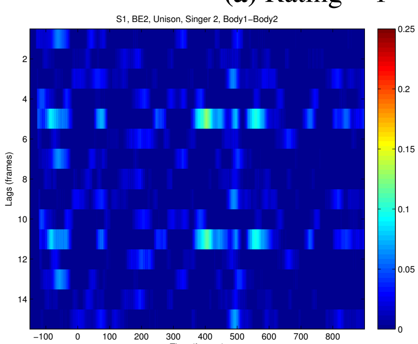
In the case of Figure 3.5, there is a band of higher transfer entropy (with some gaps) at 8 frames lag. Across conditions, there were sometimes bands of high transfer entropy at specific lags, or peaks around specific time windows. In others, any peaks were very weak, and diffused across all lags, rather than concentrated in a band at one or more particular lags. For ease of interpretation of the transfer entropy analysis results, I broke the heat maps for each signal pair and token type into one of seven categories, each with an assigned value, as shown in Figure 3.6. A rating of one (Figure 3.6a) is weak; with no values above 0.1. Two (Figure 3.6b) is more diffuse, with peaks around 25 msec in duration, between approximately 0.1 and 0.15. Three (Figure 3.6c) shows peaks around 50 msec in duration, between approximately 0.1 and 0.15, concentrated into one or more bands. Four (Figure 3.6d) shows values between approximately 0.1 and 0.15, concentrated in one or more bands, lasting 100 msec or more. Five (Figure 3.6e) shows values at or above 0.1, concentrated in one or more bands, lasting 400 msec or more. Six (Figure 3.6f) shows values consistently at or above 0.1, with some peaks above 0.15, lasting 700 msec or more. 0 indicates no tokens for that particular duet.



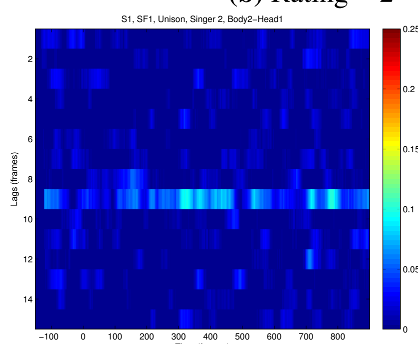
(a) Rating = 1



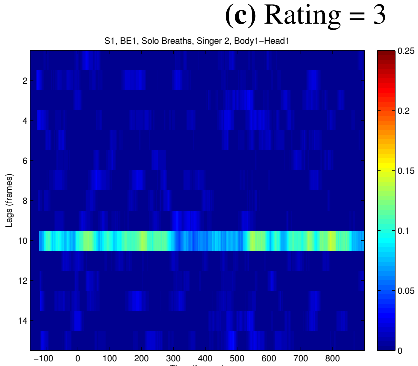
(b) Rating = 2



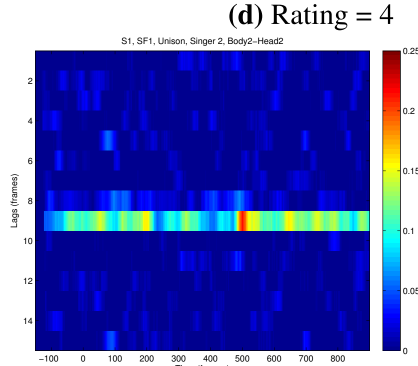
(c) Rating = 3



(d) Rating = 4



(e) Rating = 5



(f) Rating = 6

Figure 3.6: Outline of the categories of descriptions of each heat map and their corresponding numeric ratings.

Once I rated the heat maps, I plotted these values for both directions of each signal pair, to identify changes in overall pattern of TE over the course of the session. These charts are presented in two columns, representing the two possible directions between signals in each signal pair. Figure 3.7 shows the changes in rating across duets for Unison tokens in Session 3. The motion signal pairs are presented in rows. The direction of transfer entropy is presented as two columns (*A to B* corresponds to Body1 to Head1 and *B to A* corresponds to Head1 to Body1, for example). Singer 1 and Singer 3 refer to the tokens taken from annotation of Singer 1 and Singer 3's movements and singing. The numeric rating is plotted vertically; consecutive duets are plotted horizontally.

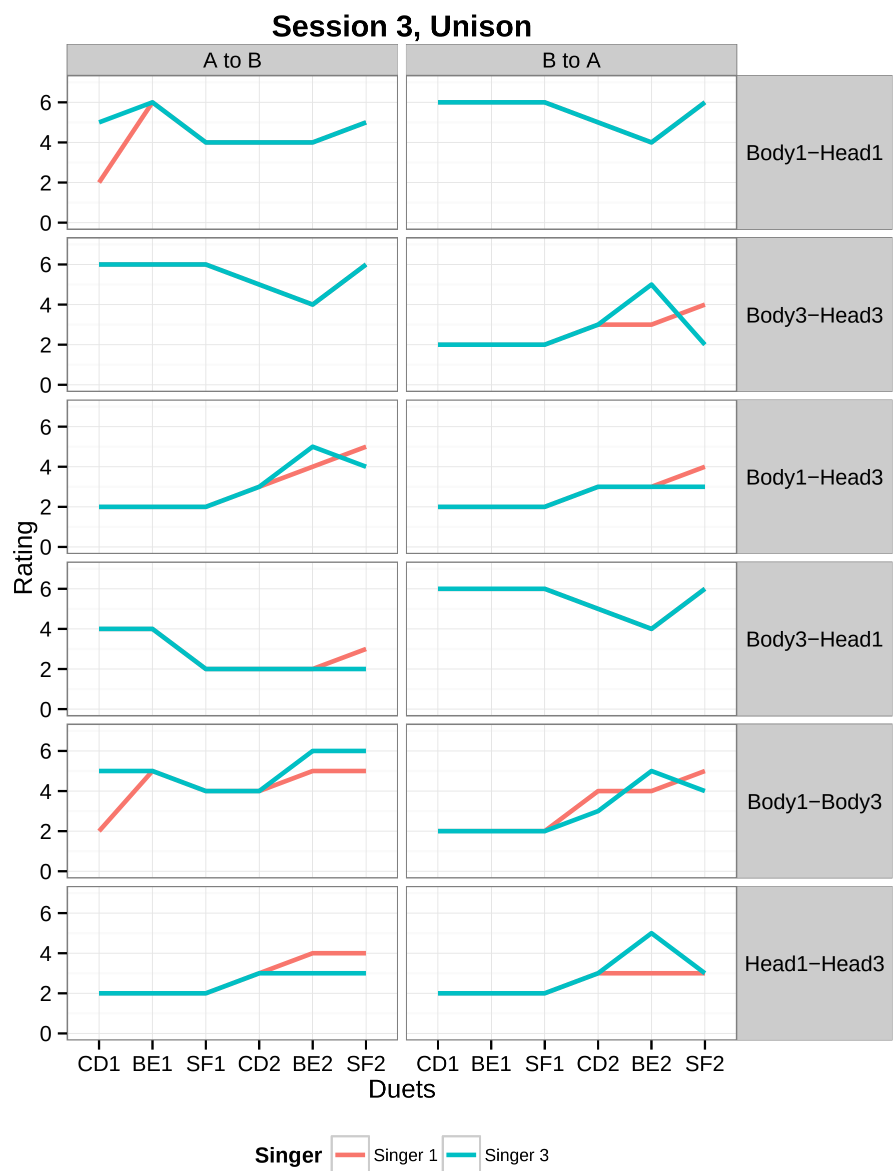
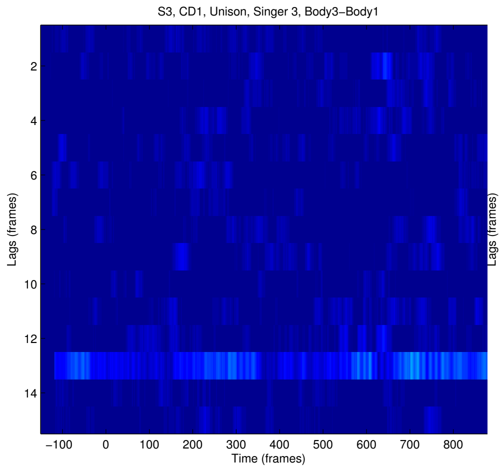
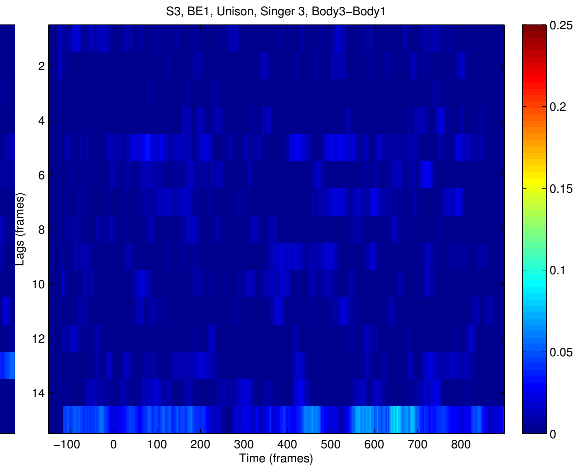


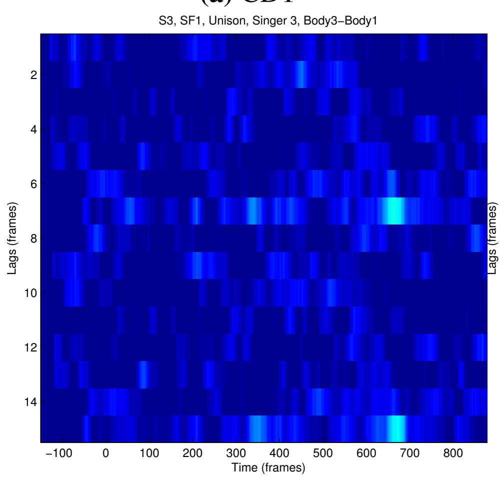
Figure 3.7: Summary of Session 3 Unison Transfer Entropy Analysis Results.



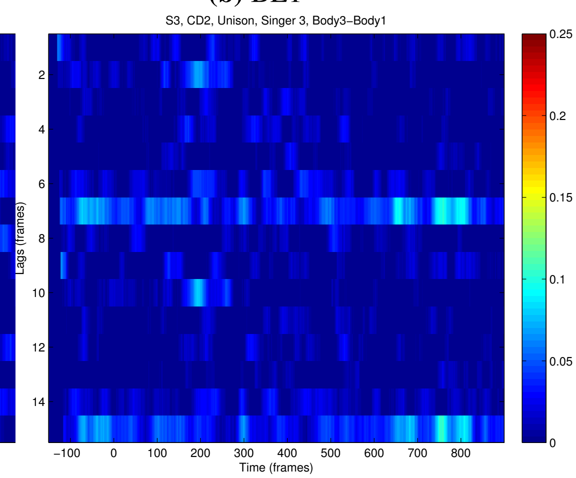
(a) CD1



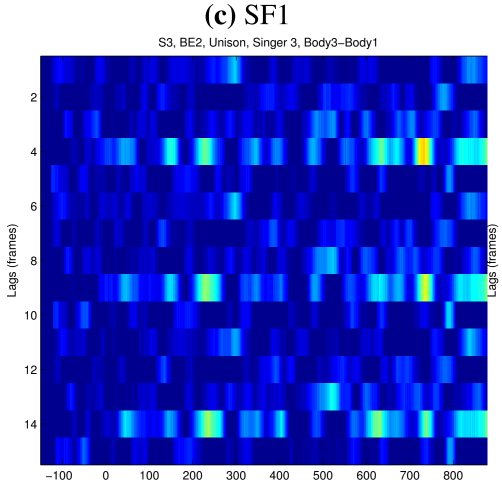
(b) BE1



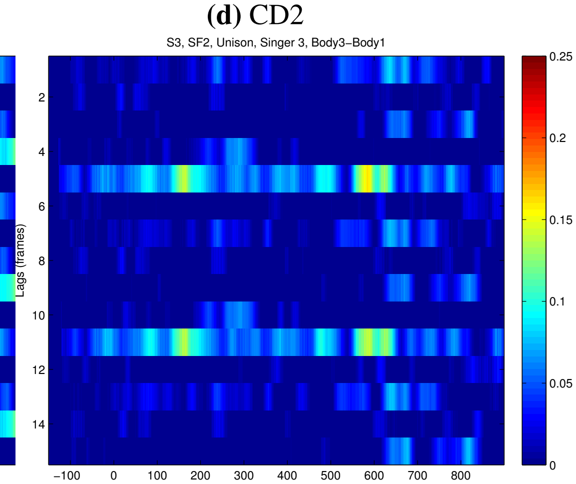
(c) SF1



(d) CD2



(e) BE2



(f) SF2

Figure 3.8: Heatmaps for Session 3, Singer 3, Body3 to Body1, Unison. Illustrates the increase in transfer entropy seen from Body 3 to Body1 in Figure 3.7.

Similar to the correlation map analysis results, there were no consistent increases in transfer entropy across the series of duets in Session 1 (see Figure C.1. There were, however, some signal pairs that showed this pattern. In Session 3, the Body1-Head3, Body1-Body3, and Head1-Head3 signal pairs showed modest increases in the level of transfer entropy over the course of the session, for both directions and both singers (see Figure 3.7.

In order to summarize the transfer entropy results further, I collapsed the between-singer motion signals into one category, and simply compared between the two possible directions. Figure 3.9 shows the results of this summary for Singer 1 in Session 1. Each column represents a direction (Singer 1 to Singer 2 and Singer 2 to Singer 1), and each row represents a token type (Unison, Solo, etc.). The number of occurrences of each rating are plotted as stacked bars, with a different color for each rating (from 4 to 6). Duets are plotted horizontally.

Figure 3.9: Session 1, Singer 1 Ratings, By Direction and Token

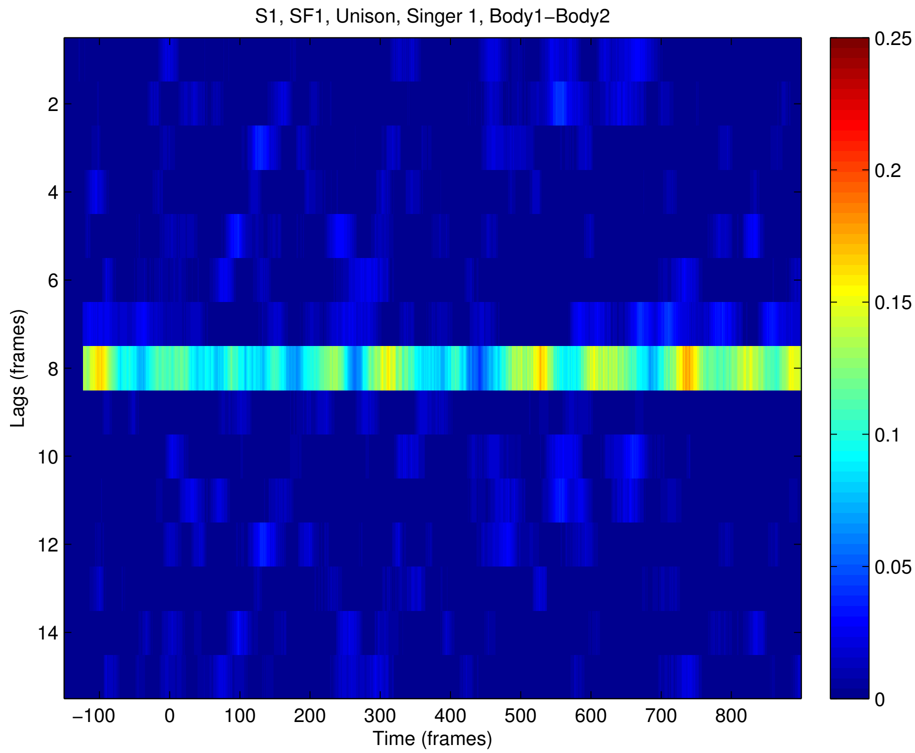
Across token types, transfer entropy did not increase in later duets. There were some asymmetries between directions, although the source of these is not apparent in the collapsed results.

Transfer entropy results were very variable across token types, sessions, and duets. One commonality is that there were not typically noteworthy peaks in transfer entropy around the token onset (0 frames). This observation includes duets and tokens where $n > 30$. Peaks often occurred later in the time window, which is 30 seconds long and therefore longer than the temporal offsets used in the correlation map analysis.

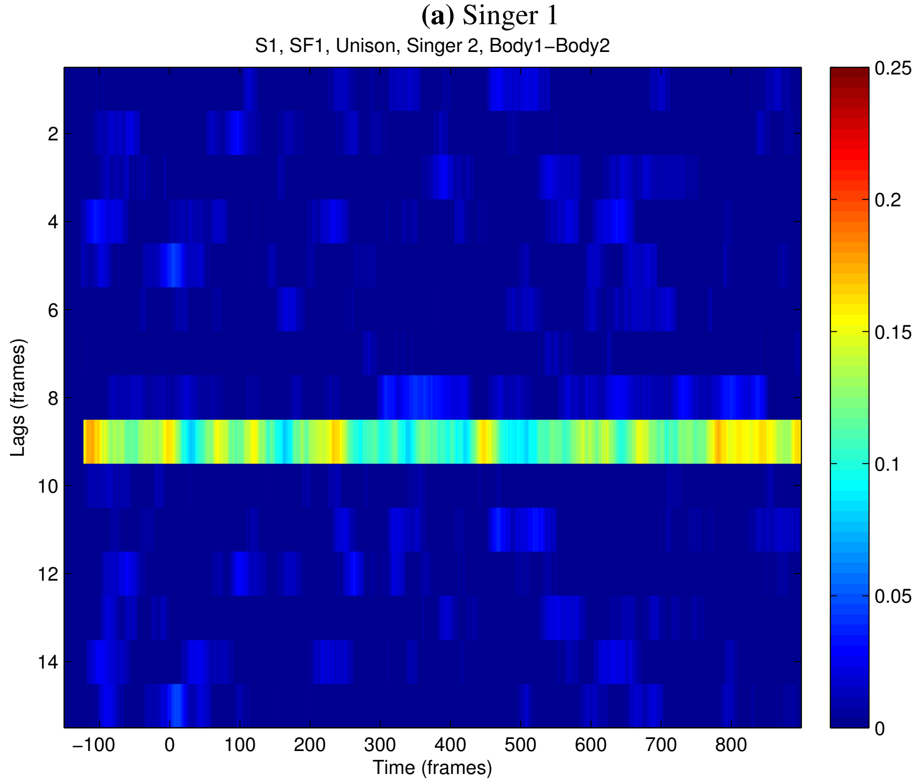
Where bands of transfer entropy did occur, the lags that showed higher, consistent levels of transfer entropy tended to repeat across signal pairs, and in both directions. Likewise, they often tended to be near to each other between singers–for instance, Singer 1 might show a band of transfer entropy at 8 frames lag, and Singer 2 might show a band at 9 frames lag (as in Figure 3.10). In the case of the unison tokens and the breaths-before-unison-sections tokens, the actual onsets of the individual tokens might only be a few frames apart. In the rest, though, the token onsets might be separated by several seconds.

There were also some asymmetries between directions for the same signal pair and duet, as in Figure 3.11, where transfer entropy was weaker for Body 2 to Body 1 than for Body 1 to Body 2, although transfer entropy was present in a band in the same frame lag.

Comparing across token types, transfer entropy was more consistently present for breaths before unison sections than breaths before solo sections, and for sections sung in unison than for solo sections. Both Unison sections and Breaths before Unison sections showed more consistent presence of transfer entropy than Looks Up or At one another.

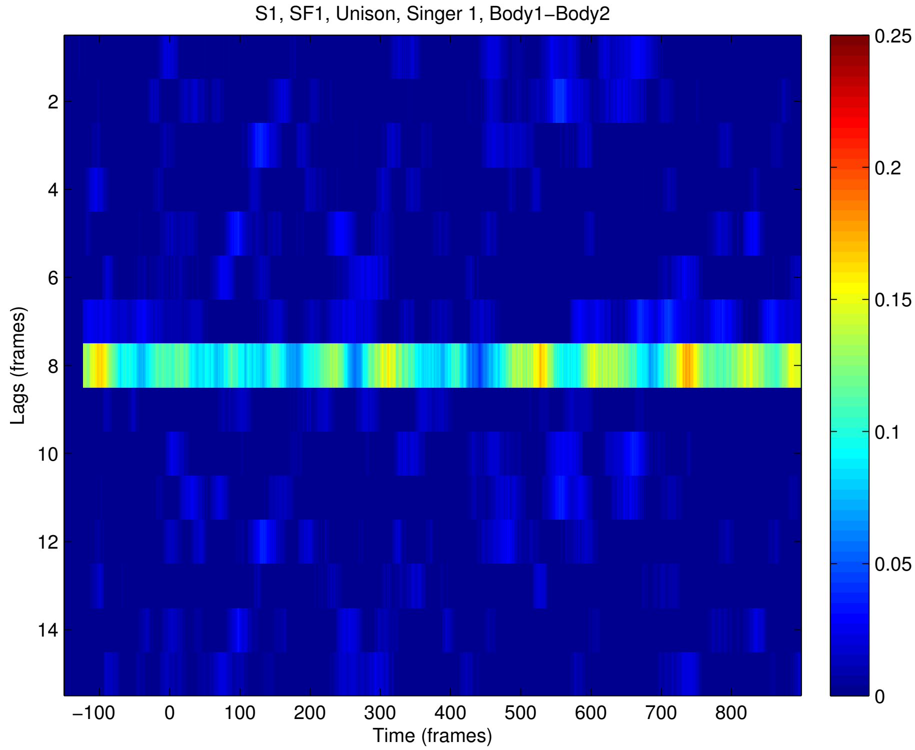


(a) Singer 1

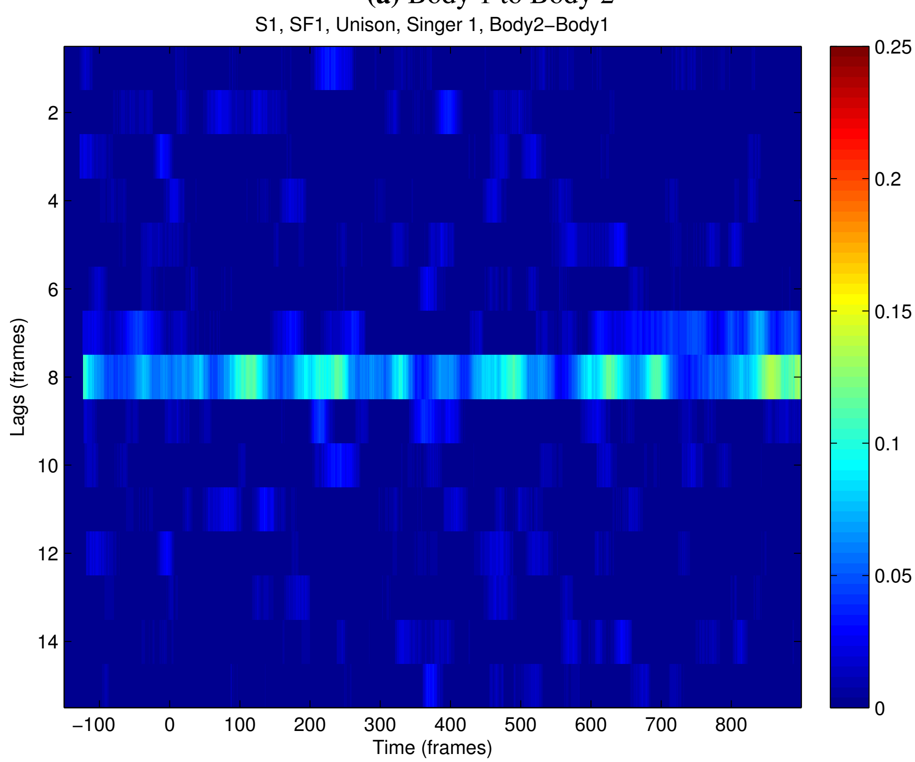


(b) Singer 2

Figure 3.10: Comparison between Singer 1 and Singer 2 for Session 1, Unison, Body1-to-Body2, SF1. A band of high transfer entropy occurs at 8 frames lag for Singer 1 and at 9 frames lag for Singer 2.



(a) Body 1 to Body 2



(b) Body 2 to Body 1

Figure 3.11: Comparison between directions for Singer 1, Scarborough Fair 1, Unison. Transfer entropy is weaker from Body 2 to Body 1 than from Body 1 to Body 2, although it largely occurs in a band at the same frame lag for both.

Chapter 4

Discussion

Williamon and Davidson [34] showed that amount that performers gesture to one another increases over time spent rehearsing and performing together. Ginsborg and King [12] showed that differences in the amount of gesture between performers is at least somewhat linked to familiarity between performers. Following these two studies, I predicted that later duets in a short recording session would show greater amounts and/or durations of glances towards one another. I also hypothesized that if gesture is used to by singers to coordinate performance, then we might see higher levels of correlation and/or transfer entropy around gestures such as eye contact and motions necessary for sound production (also referred to as ancillary gestures) such as breaths.

In order to test this, I followed Williamon and Davidson [34]'s method of annotating and quantifying the number and duration of glances between singers for six duets in a single recording session (three duets sung twice). However, my singers did not look at each other or up more often or longer during the later duets. This is counter to predictions made following Williamon and Davidson [34], where pianists' eye contact increased significantly over the course of the practice sessions. This may be due to the structure of the recording sessions in this study, where two people performed three songs through twice, taking about 20 minutes. In general, the singers did not have their music memorized, and often needed to refer to their sheet music. They had much less time to become familiar with their singing partner than the two pianists in Williamon and Davidson [34]. It may be that the kind of familiarity with both music and coperformer needed to show significant communicative gesturing is developed over a longer time span. It is also possible that the musicians' objectives in the current study differed from the objectives in Williamon and Davidson [34], where the pianists were preparing for a performance.

In addition to testing predictions on the number and duration of gestures, I also applied two quantitative analyses to motion signals extracted from the video (using optical flow analysis), as well as the rms amplitude of the audio signals. The first analysis, correlation map analysis, examines the levels of correlation between two signals over a range of temporal offsets. It is useful for establishing whether two signals are correlated positively or negatively, and whether they are synchronized (with a temporal offset of 0 seconds) or are correlated at a temporal offset (one event is related to another, but they do not necessarily coincide). The second analysis, transfer entropy analysis, calculates the amount of information transferred from one signal to another. While it does not show that events in one signal are directly causing the events in another, it does show how predictable future events in one signal are based on past events in both signals.

Transfer entropy analysis relies on high sample sizes ($n \geq 30$) for statistical reliability, meaning that the experimenter must isolate individual tokens of each condition being tested over the course of the experiment. In the present study, I pulled out time windows surrounding the tokens (glances, two different types of breaths, and the onsets of solo versus unison phrases) identified during annotation. Because of this, sample sizes in the present study were insufficient for producing statistically significant results. Correlation map analysis, meanwhile, can provide a picture of changes in correlation between two signals over the course of an entire recording. However, in the present study, with the use of a non-causal filter, it does not capture any directional changes in that relationship. Transfer entropy analysis, because it only examines past samples, is inherently causal (in the sense of Wiener [33]), and as such can illustrate asymmetries between the two signals.

In the present study, I predicted that the level of correlation between singers would increase over the course of each recording session, following the increase in number and duration of gestures I predicted following Williamon and Davidson. In general, however, the correlation map analysis did not show significant increases in the amount of correlation between singers over the course of either Session 1 or 3, even at larger offset windows. The number and duration of glances up did not increase over the course of each recording session either, so technically the two did not pattern dissimilarly. Similarly to the results from the qualitative analysis, the lack of increase in correlation over time could be because the current study does not look at a long enough time span of coperformer interaction. It could also be that increased levels of gestures such as eye contact and hand gestures simply do not translate to higher levels of correlation between performers' movements.

Similar to the results of the correlation map analysis, the results of the transfer entropy analysis did not show significantly higher levels of transfer entropy for later duets in each recording session. If there is any link between higher levels of information transfer and particular movements or gestures, it is not reflected in the frames immediately before or after the onset of the token–for instance, before the beginning of a phrase sung in unison. There are two possibilities–that gestures by one singer do not have an immediate effect on the gestures of the other singer, or that gestures between singers show greater levels of transfer entropy at higher temporal lags than were examined in this study. This would be analogous to the wider temporal offsets examined in the correlation map analysis.

The transfer entropy analysis was severely limited by sample size. It is possible that altering the current study to examine several recording sessions, rather than several duets within one recording session, would yield sufficient sample sizes to make claims about the statistical significance of the transfer entropy results. However, samples sizes did equal or exceed 30 for a token types and duets, and those did not necessarily show stronger transfer entropy around 0 frames.

There were some differences in transfer entropy between token types. Unison sections showed higher levels of transfer entropy than solo sections, and breaths before unison sections showed higher levels of transfer entropy than breaths before solo sections. Looks up and looks at one another both showed lower levels of transfer entropy than the rest of the token types. So, it does appear that there is some relationship between musical structure and singer movements. This is in keeping with the prediction that movements during unison sections might pattern differently than movements during solo sections.

In many ways, the present study fails to attain significant results. In the case of the correlation map analysis, there was no comparison between types of glances, and although there was some examination of sections performed in unison, they were only compared with correlations results for the entire duet, not sections sung in unison, leaving the examination of whether levels of correlation correspond to different types of musical structure or gesture incomplete.

As mentioned earlier, the singers relied heavily on sheet music during all recording sessions. This may have had some influence on the amount of gestures between them compared to Williamon and Davidson's pianists, as they were likely not as experienced with sight reading, and therefore likely less confident looking up from their music. Notably, during the two duets in Session 2 where the singers did not use sheet music, they spent significantly more time looking at each other than during the duets where they held sheet music in the same session. Future studies might benefit from requiring that their participants memorize the pieces being performed.

The structure of this study also differed from the study in Williamon and Davidson, in that the singers only met for one approximately thirty-minute recording session, rather than several recording sessions over several days. Musicians' gestures may not increase (and that levels of correlation between them do not increase) over such a short span of time. A useful extension of the present study would compare these results with singers who have worked together over several rehearsals, more closely following Williamon and Davidson [34] and Ginsborg and King [12].

Chapter 5

Conclusion

The present study examined the gestures of pairs of singers over the course of six duets to establish whether (and how) singers' gestures change over the time, as well as whether correlation between singers movements increases over time. In particular, the study used optical flow analysis, correlation map analysis and transfer entropy analysis to analyze singers' movements and singing as related time series. I annotated audio and video recordings of singers to isolate times when singers sang alone or in unison, times when they breathed before solo and unison phrases, times when they looked at one another, and times when they looked up from their sheet music. Analysis of the number and duration of times when the singers looked up or at each other showed that singers did not look at each other more during later duets, suggesting that use of eye contact to communicate may be limited by unfamiliarity with the pieces being performed, or by unfamiliarity with the coperformer.

Using optical flow analysis, I isolated four regions of interest for video of the singers, reducing the movement of each singer's head and body to 5 time series. These signals, along with the rms amplitude of the audio recordings, formed the basis of the correlation map analysis and transfer entropy analysis.

I generated correlation maps for fifteen different signal pairs. In general, the correlation maps showed that while signal pairs did have somewhat predictable patterns of correlation across duets, levels of correlation themselves did not increase over the course of the recording session.

Using transfer entropy analysis, I examined relationships between six signal pairs for particular gestures: looks at one another, looks up, unison phrases, solo phrases, breaths before unison phrases, and breaths before solo phrases. In general, transfer entropy was not stronger during later duets than earlier duets. Furthermore, there were no immediate peaks in transfer entropy following onset of the gesture in being examined, suggesting that influence of one signal on another is not instantaneous. Furthermore, looking at levels and patterns of transfer entropy over the course of entire recording sessions, it seems that the relationship between signals does not progress in a predictable fashion. This follows the results of the correlation map analysis.

The duration of looks, levels of correlation, and level of transfer entropy did not show significant, if any, increase over the duration of the recording sessions examined. There are two possibilities for the lack of increase. The first is that nonverbal communication and correlation between singers' movements do not increase over time in performance. The second possibility is that nonverbal communication does increase, but at a slower rate that was not captured within the duration of the short performances examined in this study. Future studies should seek highly skilled singers, recruit them to sing together during several sessions times over a longer period, and use larger offset windows (in the case of correlation map analysis) and lags (in the case of the transfer entropy analysis) in order to answer the question of whether nonverbal communication and correlation levels do not increase, or if they merely increase slowly.

Bibliography

[1] ELAN. Retrieved December 4, 2015, from https://tla.mpi.nl/tools/tla-tools/elan/. Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands., 2015. → pages 8

[2] A. V. Barbosa, H. C. Yehia, and E. Vatikiotis-Bateson. Matlab toolbox for audiovisual speech processing. In *Proceedings of the International Conference on Auditory-Visual Speech Processing*, 2007. → pages 12

[3] A. V. Barbosa, H. C. Yehia, and E. Vatikiotis-bateson. Linguistically Valid Movement Behavior Measured Non-Invasively. In *Auditory Visual Speech Processing*, pages 173–177, 2008. → pages 4, 18

[4] A. V. Barbosa, R.-M. Déchaine, E. Vatikiotis-Bateson, and H. C. Yehia. Quantifying time-varying coordination of multimodal speech signals using correlation map analysis. *The Journal of the Acoustical Society of America*, 131(3):2162–72, Mar. 2012. ISSN 1520-8524. doi:10.1121/1.3682040. URL http://www.ncbi.nlm.nih.gov/pubmed/22423712. → pages 4, 12

[5] N. Bedo, U. Ribary, and L. M. Ward. Fast dynamics of cortical functional and effective connectivity during word reading. *PloS one*, 9(2):e88940, Jan. 2014. ISSN 1932-6203. doi:10.1371/journal.pone.0088940. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3925174&tool=pmcentrez&rendertype=abstract. → pages 4

[6] E. Bigand, R. Parncutt, and F. Lerdahl. Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, 58(1):125–141, 1996. → pages 2

[7] P. Boersma. Praat, a system for doing phonetics by computer. *Glot International*, 5 (9/10):341–345, 2001. → pages 9

[8] G. Chew. Articulation and phrasing, 2015. URL http://www.oxfordmusiconline.com.ezproxy.library.ubc.ca/subscriber/article/grove/music/40952. → pages 2

[9] H. A. M. Clayton. *Coordination Between Players in Musical Performance*. PhD thesis, University of Edinburgh, 1985. → pages 1

[10] R. Cooper, O. Harrison, J. Kamen, and B. Roylance. Concert for george [digital]. United Kingdom: Warner Home Video, 2003. → pages 2

[11] C. A. Fowler, M. J. Richardson, K. L. Marsh, and K. D. Shockley. Language Use, Coordination, and the Emergence of Cooperative Action. In A. Fuchs and V. K. Jirsa, editors, *Coordination: Neural, Behavioral and Social Dynamics: Understanding Complex Systems*, pages 261–279. 2008. → pages 1

[12] J. Ginsborg and E. King. Gestures and glances : The Effects of Familiarity and Expertise on Singers and Pianists Bodily Movements in Ensemble Rehearsals. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, and P.-S. Eerola, editors, *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music*, number Escom, pages 159–164, Jyväskylä, Finland, 2009. → pages 3, 4, 5, 30, 32

[13] W. Goebl and C. Palmer. Synchronization of Timing and Motion Among Performing Musicians. *Music Perception*, 26(5):427–439, 2009. → pages 1

[14] G. Gomez-Herrero, W. Wu, K. Rutanen, M. C. Soriano, G. Pipa, and R. Vicente. Assessing coupling dynamics from an ensemble of time series. Aug. 2010. URL http://arxiv.org/abs/1008.0539. → pages 14

[15] B. K. Horn and B. G. Schunck. Determining Optical Flow. *Artificial Intelligence*, 17 (1-3):185–203, Aug. 1981. ISSN 00043702. doi:10.1016/0004-3702(81)90024-2. URL http://linkinghub.elsevier.com/retrieve/pii/0004370281900242. → pages 4

[16] D. B. Huron. *Sweet Anticipation: Music and the Psychology of Expectation*. MITPress, 2006. → pages 2

[17] S. Kawase. Gazing behavior and coordination during piano duo performance. *Attention, Perception & Psychophysics*, 76(2):527–40, Feb. 2014. ISSN 1943-393X. doi:10.3758/s13414-013-0568-0. URL http://www.ncbi.nlm.nih.gov/pubmed/24170378. → pages 2

[18] P. E. Keller. Attentional Resource Allocation in Musical Ensemble Performance. *Psychology of Music*, 29(1):20–38, Apr. 2001. ISSN 0305-7356. doi:10.1177/0305735601291003. URL http://pom.sagepub.com/cgi/doi/10.1177/0305735601291003. → pages 1, 2

[19] P. E. Keller. Joint Action in Music Performance. In F. Morganti, A. Carassa, and G. Riva, editors, *Enacting Intersubjectivity: A Cognitive and Social Perspective on the Study of Interactions*, number May, pages 205–221. 2008. → pages 1

[20] H. Lausberg and H. Sloetjes. Coding gestural behavior with the NEUROGES–ELAN system. *Behavior research methods*, 41:841–849, 2009. ISSN 1554-351X. doi:10.3758/BRM.41.3.841. → pages 8

[21] F. Lerdahl. *Tonal Pitch Space*. Oxford University Press, 2001. → pages 2

[22] F. Lerdahl and C. L. Krumhansl. Modeling Tonal Tension. *Music Perception: An Interdisciplinary Journal*, 24(4):329–366, 2007. → pages 2

[23] J. Page and R. Plant. The Battle of Evermore. On Led Zeppelin IV [Digital]. United Kingdom: Atlantic, 1971. → pages 7

[24] J. Phillips and M. Phillips. California Dreamin' [Recorded by The Mamas and the Papas]. On If You Can Believe Your Eyes and EArs [Digital]. United States: Dunhill Records, 1965. → pages 7

[25] T. Schreiber. Measuring information transfer. *Physical review letters*, 85(2):461–4, July 2000. ISSN 1079-7114. doi:10.1103/PhysRevLett.85.461. URL http://www.ncbi.nlm.nih.gov/pubmed/10991308. → pages 4

[26] F. Seddon and M. Biasutti. A comparison of modes of communication between members of a string quartet and a jazz sextet. *Psychology of Music*, 37(4):395–415, June 2009. ISSN 0305-7356. doi:10.1177/0305735608100375. URL http://pom.sagepub.com/cgi/doi/10.1177/0305735608100375. → pages 2

[27] K. Shockley, M.-V. Santana, and C. a. Fowler. Mutual Interpersonal Postural Constraints are Involved in Cooperative Conversation. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2):326–332, 2003. ISSN 0096-1523. doi:10.1037/0096-1523.29.2.326. URL http://doi.apa.org/getdoi.cfm?doi=10.1037/0096-1523.29.2.326. → pages 1

[28] Traditional and P. Simon. Scarborough Fair/Canticle [Recorded by Paul Simon and Art Garfunkel]. On Parsley, Sage, Rosemary, and Thyme [Digital]. United States: Columbia, 1966. → pages 7

[29] E. Vatikiotis-Bateson, A. V. Barbosa, and C. T. Best. Articulatory coordination of two vocal tracts. *Journal of Phonetics*, 44:167–181, May 2014. ISSN 00954470. doi:10.1016/j.wocn.2013.12.001. URL http://linkinghub.elsevier.com/retrieve/pii/S0095447013000922. → pages 4, 12

[30] B. W. Vines, M. M. Wanderley, C. L. Krumhansl, R. L. Nuzzo, and D. J. Levitin. *Performance Gestures of Musicians : What Structural and Emotional Information Do They Convey ?* Springer-Verlag, 2004. → pages 1

[31] B. W. Vines, C. L. Krumhansl, M. M. Wanderley, and D. J. Levitin. Cross-modal interactions in the perception of musical performance. *Cognition*, 101(1):80–113, Aug. 2006. ISSN 0010-0277. doi:10.1016/j.cognition.2005.09.003. URL http://www.ncbi.nlm.nih.gov/pubmed/16289067. → pages 2, 3

[32] M. M. Wanderley. Quantitative Analysis of Non-obvious Performer Gestures. In I. Wachsmuth and T. Sowa, editors, *Gesture and Sign Language in Human-Computer Interaction Lecture Notes in Computer Science*, pages 241–253. Springer-Verlag, 2002. → pages 1, 8

[33] N. Wiener. The theory of prediction. *Modern mathematics for engineers*, 1:125–139, 1956. → pages 4, 31

[34] A. Williamon and J. W. Davidson. Williamon&Davidson2002.pdf. *Musicae Scientiae*, V1(1):53–72, 2002. → pages 1, 3, 4, 5, 8, 15, 30, 32

Appendix A

Questionnaire

The following page contains the questionnaire I used to establish the duration of subjects' musical training, their primary instrument, as well as their impressions of their performance during the recording session–for example, points where they made mistakes and points where they felt they did well.

Before the study:

How many years have you been singing in a controlled environment (i.e. lessons, focused individual practice, etc.)?____

How many years of vocal ensemble rehearsal and performance experience do you have? At what level (i.e., high school, collegiate, in bands, etc.)______________________

Have you sung with your partner in any ensembles of less than approximately 20 people?____

Do you play any other instruments? If so, what instruments?________________________
____
____

How many years have you played this/these instrument/s?_______________________

How many years of focused lessons or practice have you received on these instrument(s)? How many years of ensemble performance (orchestra, small ensemble, band, etc.)_______
____
____

Would you describe voice as your primary instrument? If not, which instrument is your primary instrument?____

Approximately how much time did you spend practicing these duets? _______________

Are there any places in the duets that you feel are more challenging than others? If so, why?____

After the study:

How do you feel about your performance of the duets?__________________________________

Were there any places in the duets that you felt that you made mistakes?_____________
____
____

Were there any places in the duets that you felt sounded particularly good? __________________
____

Did you and your partner use any particular types of nonverbal communication during your performance? If so, what were they? _____________________________________
____

Other comments:
____
____
____
____
____

Appendix B

Correlation Map Analysis Results

This appendix contains correlation map analysis results referred to in Section 3.2. The results are categorized by type of signal pair (Within-Singer pairs, Between-Singer Motion Signal Pairs, and Between-Singer rms amplitude pairs), by Unison sections, and by size of offset. Table B.1 shows an abbreviated example of the total number of samples in each duet in session 1, the number of hits (correlation values higher than 0.5, and number of hits as a percentage of the total number of samples, from -0.25 to 0 seconds offset.

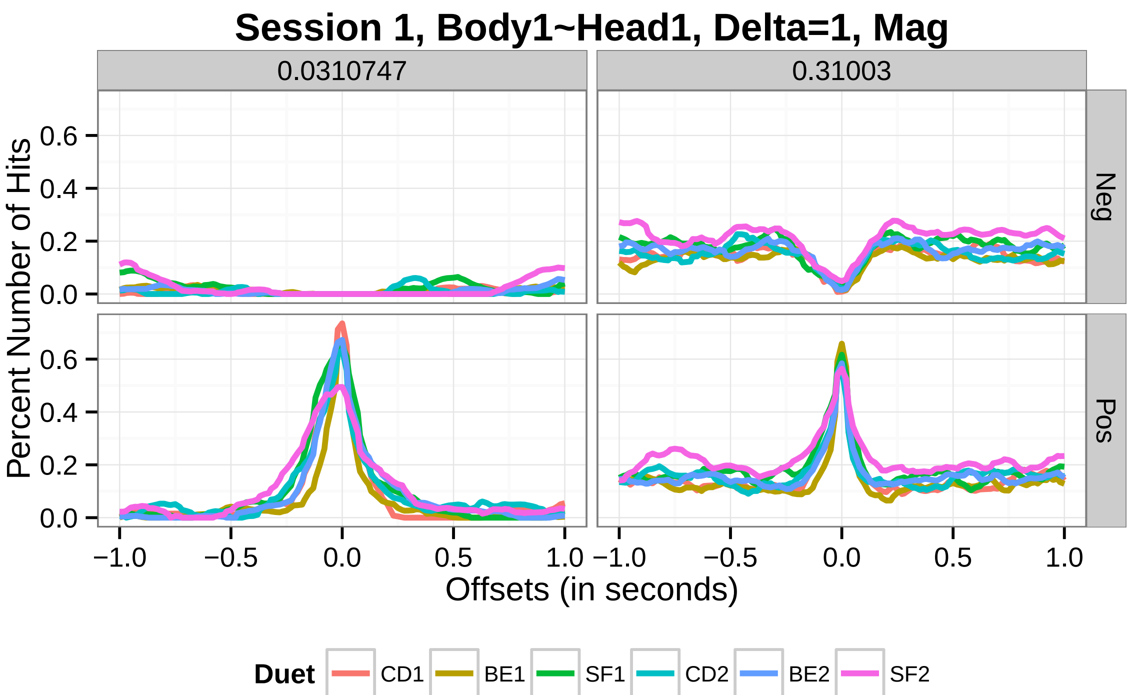
Table B.1: Number of Hits and Percent Number of Hits

Duets	CD1		BE1		SF1		CD2		BE2		SF2	
Total Samples	10912		14580		9154		10804		13108		9038	
Offset (seconds)	Hits	% Hits	Hits	% Hits	Hits	% Hits	Hits	% Hits	Hits	% Hits	Hits	% Hits
-0.25	797	0.07	1421	0.10	1086	0.12	975	0.09	1512	0.12	1061	0.12
-0.23	976	0.09	1565	0.11	989	0.11	908	0.08	1414	0.11	1035	0.11
-0.22	1147	0.11	1490	0.10	920	0.10	737	0.07	1491	0.11	1071	0.12
-0.2	1228	0.11	1491	0.10	1033	0.11	902	0.08	1546	0.12	1122	0.12
-0.18	1100	0.10	1316	0.09	1003	0.11	1079	0.10	1589	0.12	1219	0.13
-0.17	1057	0.10	1540	0.11	1092	0.12	1060	0.10	1485	0.11	1294	0.14
-0.15	1011	0.09	1466	0.10	958	0.10	974	0.09	1537	0.12	1099	0.12
-0.13	1171	0.11	1515	0.10	902	0.10	1054	0.10	1417	0.11	1004	0.11
-0.12	1042	0.10	1221	0.08	787	0.09	963	0.09	1231	0.09	957	0.11
-0.1	981	0.09	1086	0.07	704	0.08	875	0.08	1172	0.09	893	0.10
-0.08	798	0.07	1008	0.07	727	0.08	711	0.07	1146	0.09	941	0.10
-0.07	903	0.08	1025	0.07	1025	0.11	895	0.08	1197	0.09	1082	0.12
-0.05	1078	0.10	1321	0.09	1213	0.13	1298	0.12	1216	0.09	1014	0.11
-0.03	683	0.06	873	0.06	789	0.09	993	0.09	861	0.07	731	0.08
-0.02	187	0.02	460	0.03	324	0.04	471	0.04	466	0.04	460	0.05
0	128	0.01	308	0.02	229	0.03	390	0.04	415	0.03	429	0.05

B.1 Whole Song Correlations

B.1.1 Within-Singer Motion Signal Pairs

In Figure B.1, Body1-Head1 shows a trough in negative correlation for η=0.31003 between -0.3 and 0.3 seconds, and a peak in positive correlation for both η=0.031074 and 0.31003 between -0.5 and 0.5 seconds. Body2-Head2 shows similar patterns of correlation, but with weaker peaks of positive correlation.



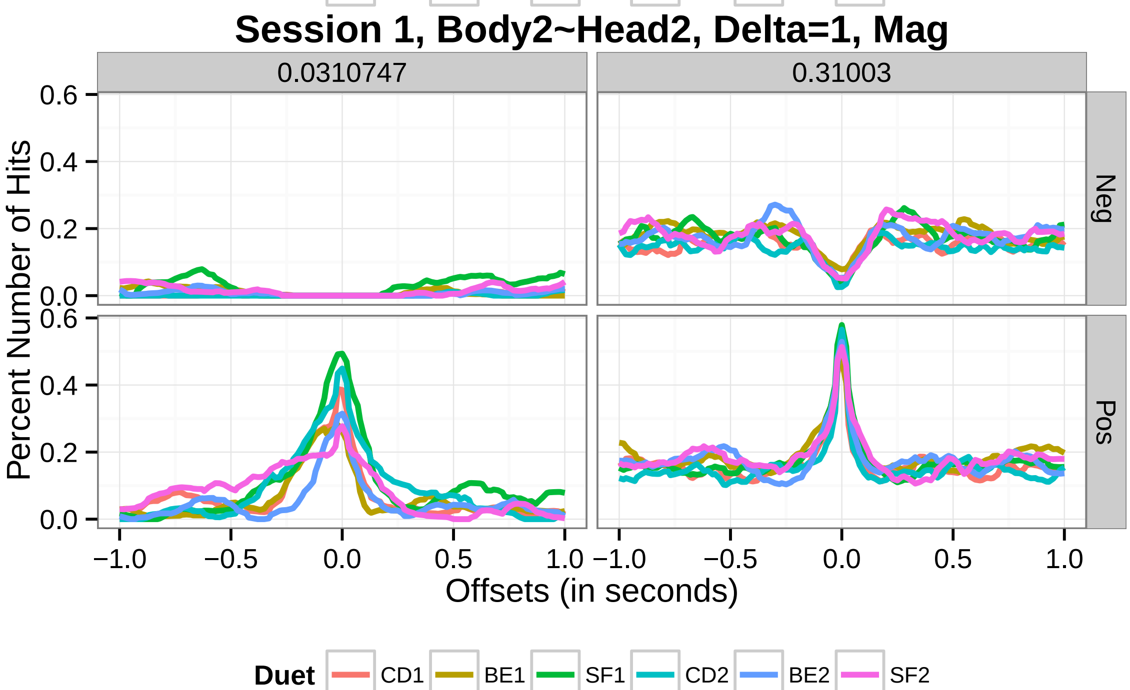


Figure B.1: Within-singer motion signal pairs for Session 1.

Figure B.2 shows the within-singer motion signal pairs for Session 3. These patterned similarly to the within-singer motion signal pairs in Session 1. Scarborough Fair 2 did appear to have larger peaks in positive correlation than earlier duets.

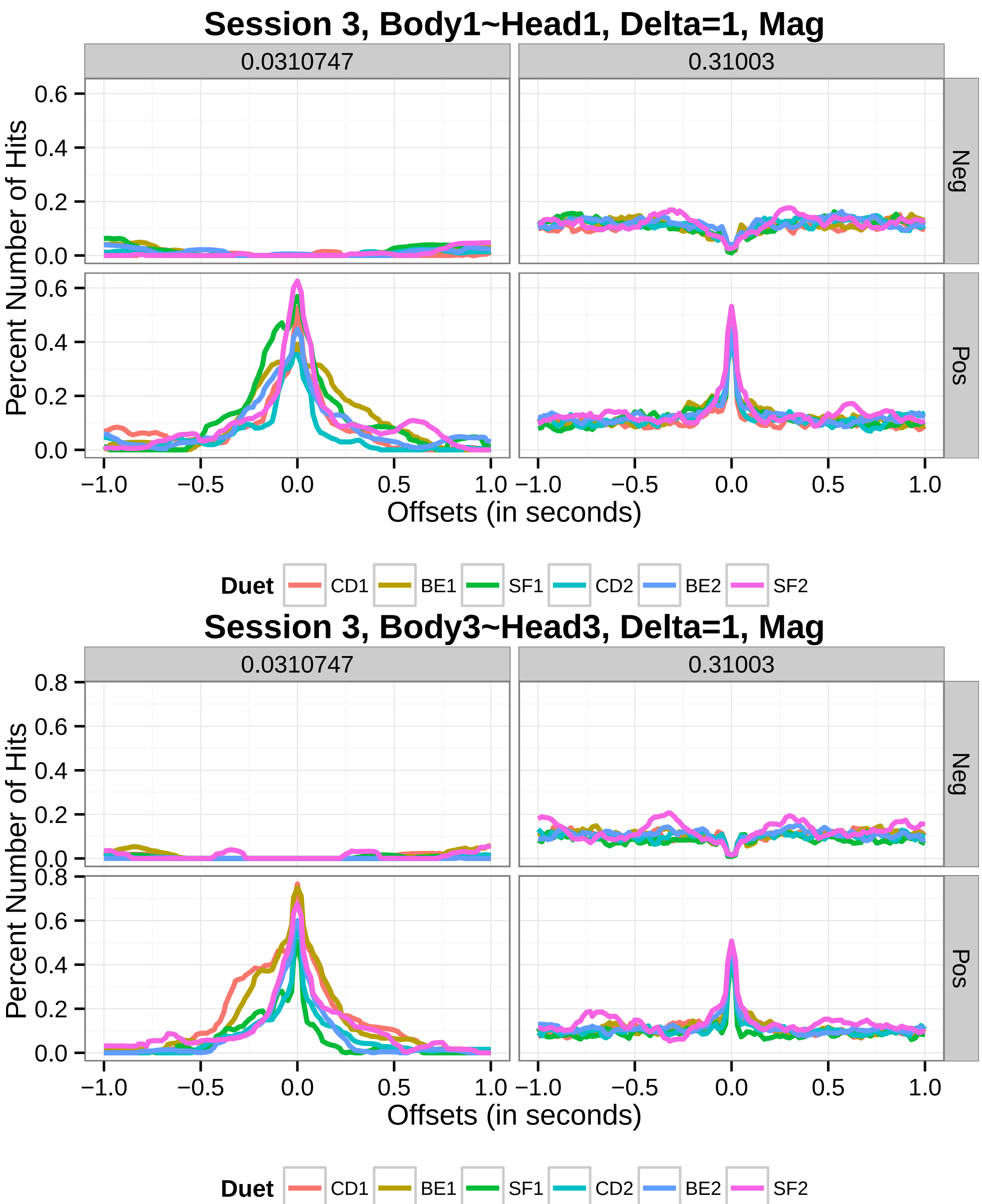


Figure B.2: Within-singer motion signal pairs for Session 3.

B.1.2 Between-Singer Motion Signal Pairs

Between-singer head signal pairs for both Session 1 and Session 3 showed similar pattern of troughs and peaks as within-singer motion signal pairs and between-singer body motion signal pairs, but with lower levels of positive correlation (less than 40%).

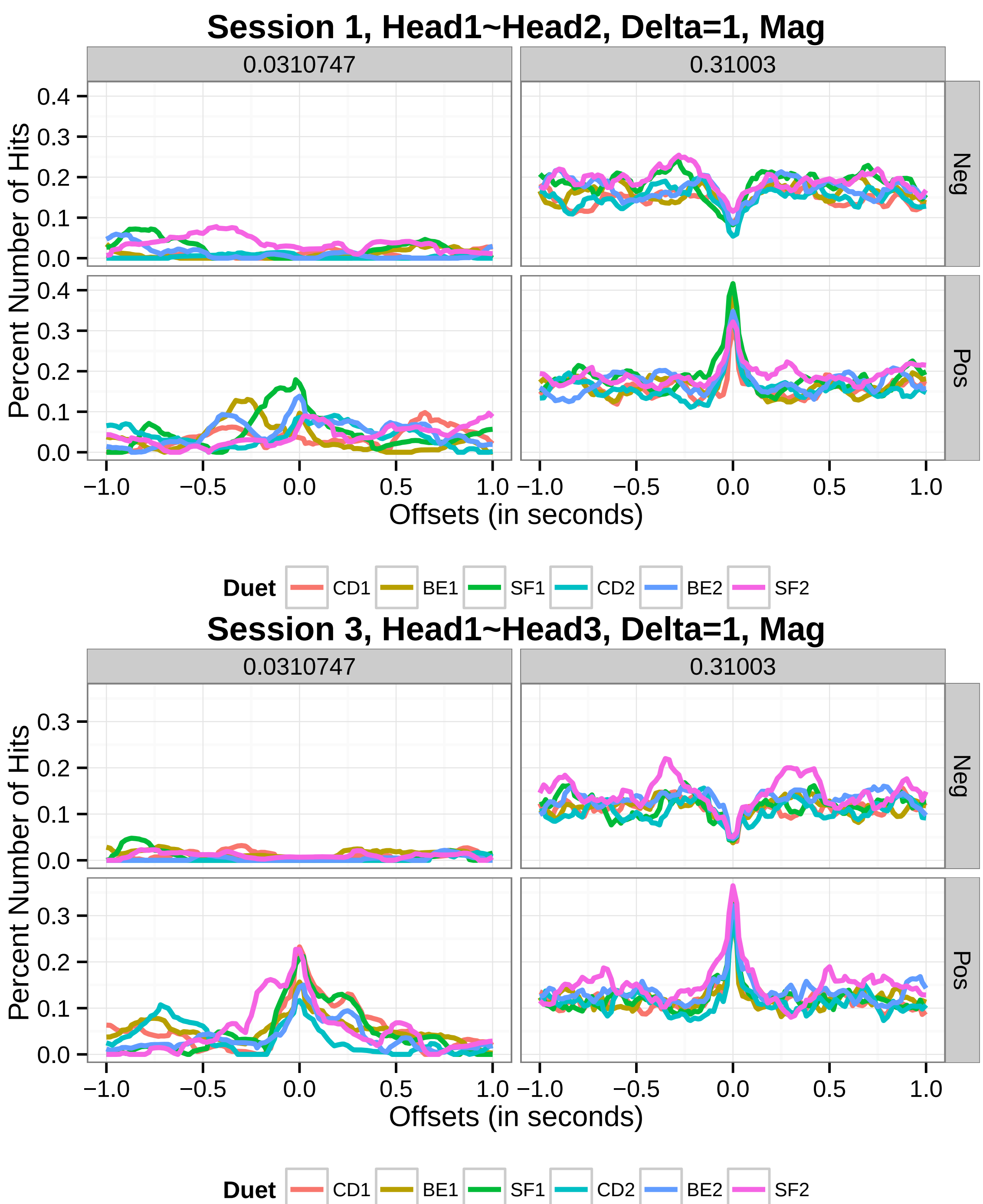
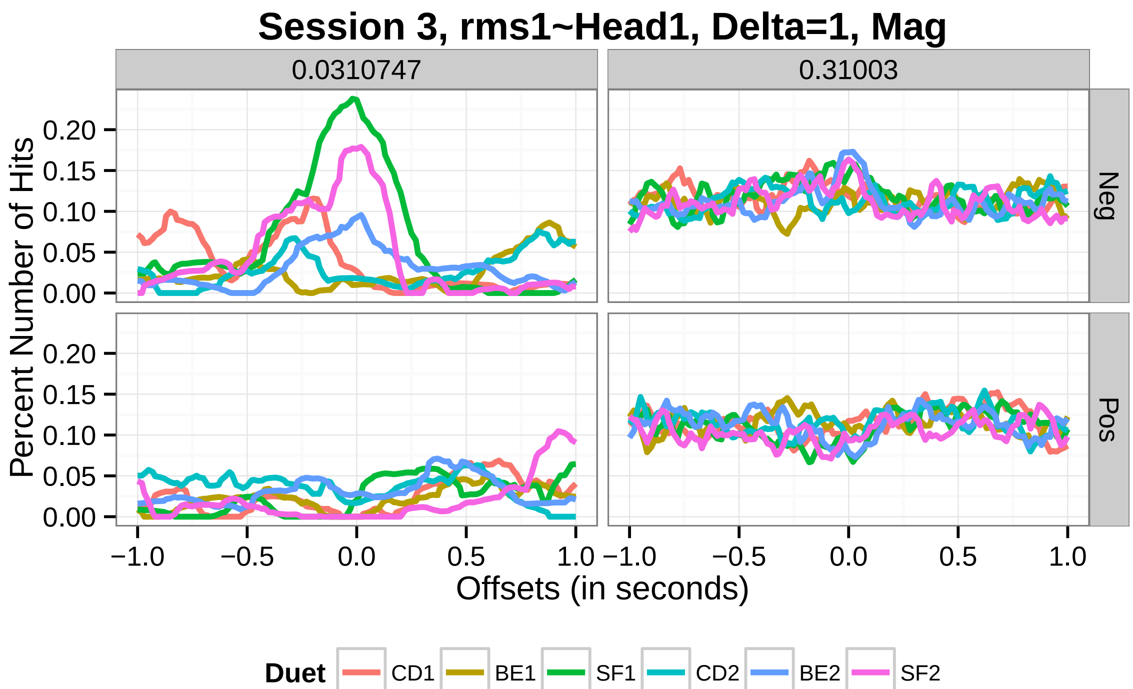


Figure B.3: Between-singer head motion signal pairs for Session 1 and Session 3

B.1.3 Within-Singer RMS Amplitude-Motion Signal Pairs

Overall, within-singer rms amplitude-motion signal pairs showed much looser and lower peaks in correlation than motion signal pairs. rms1∼Head1 in Session 3 showed some small peaks (below 30%) in negative correlation for Scarborough Fair 1 and 2. rms3∼Head3 showed peaks in negative correlations around 0 for Battle of Evermore 1 and 2 and Scarborough Fair 2.



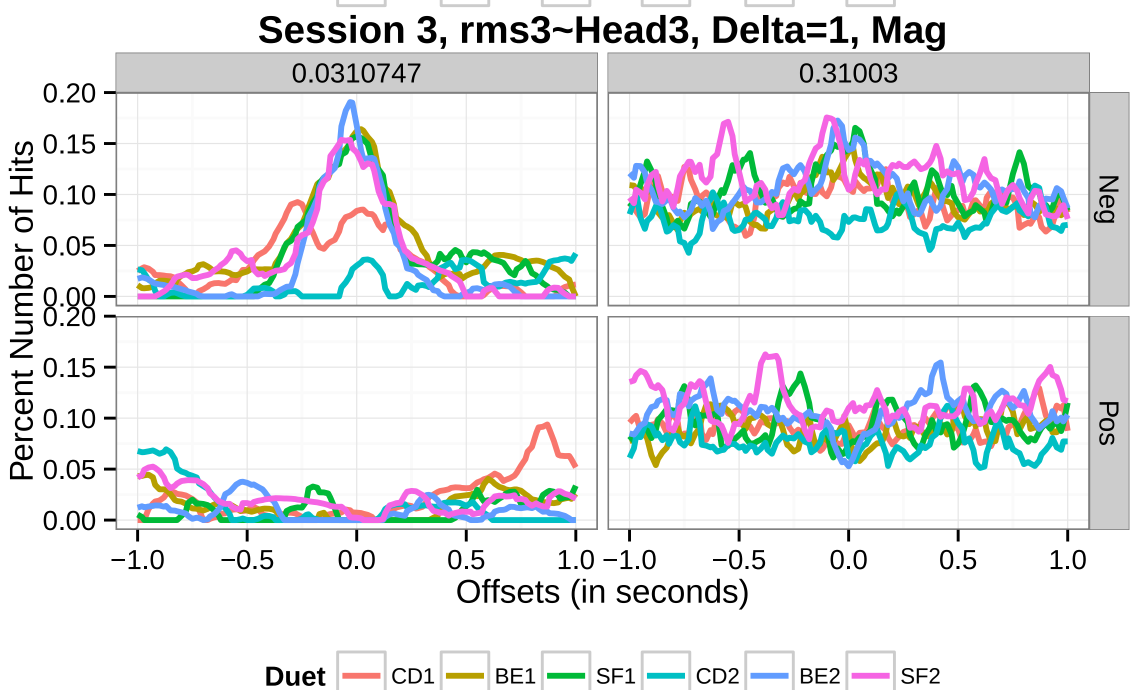
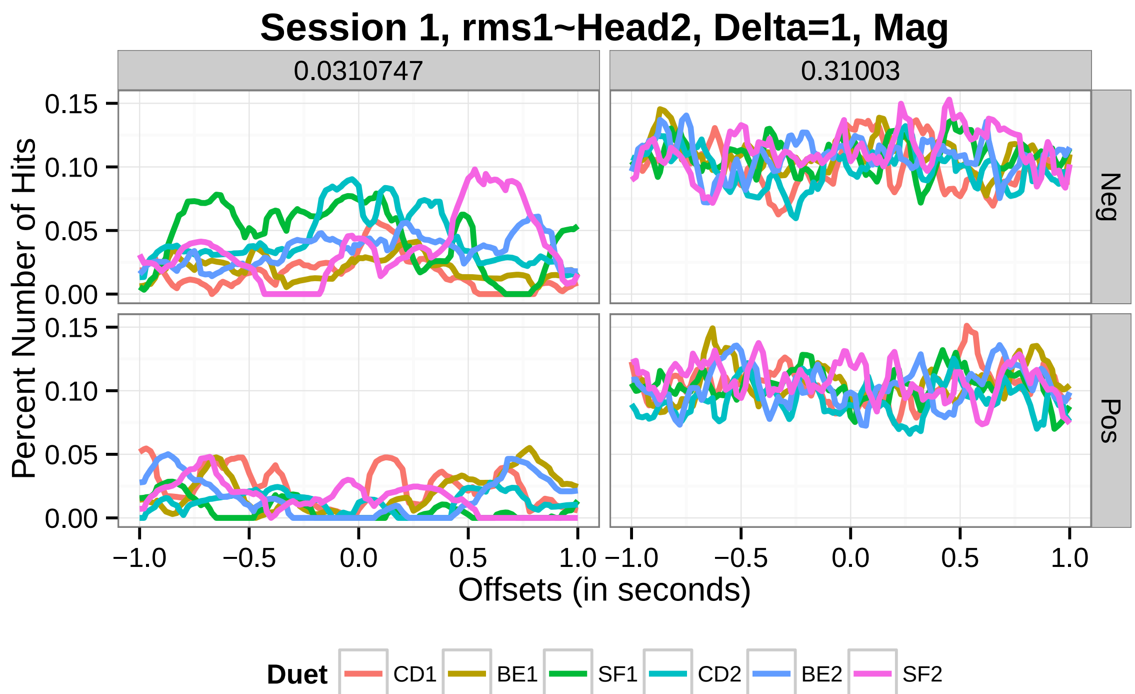


Figure B.4: Session 3 within-singer rms amplitude/head motion signal pairs.

B.1.4 Between-Singer RMS Amplitude-Motion Signal Pairs

Between-singer rms amplitude-motion signal pairs also showed loose, low peaks, which did not cluster at any particular offset. Figure B.5 shows an example of the rms amplitude-head motion signal pairs in Session 1, where correlation was consistently lower than 15% for all duets. Correlation was not generally higher for later duets in the session.



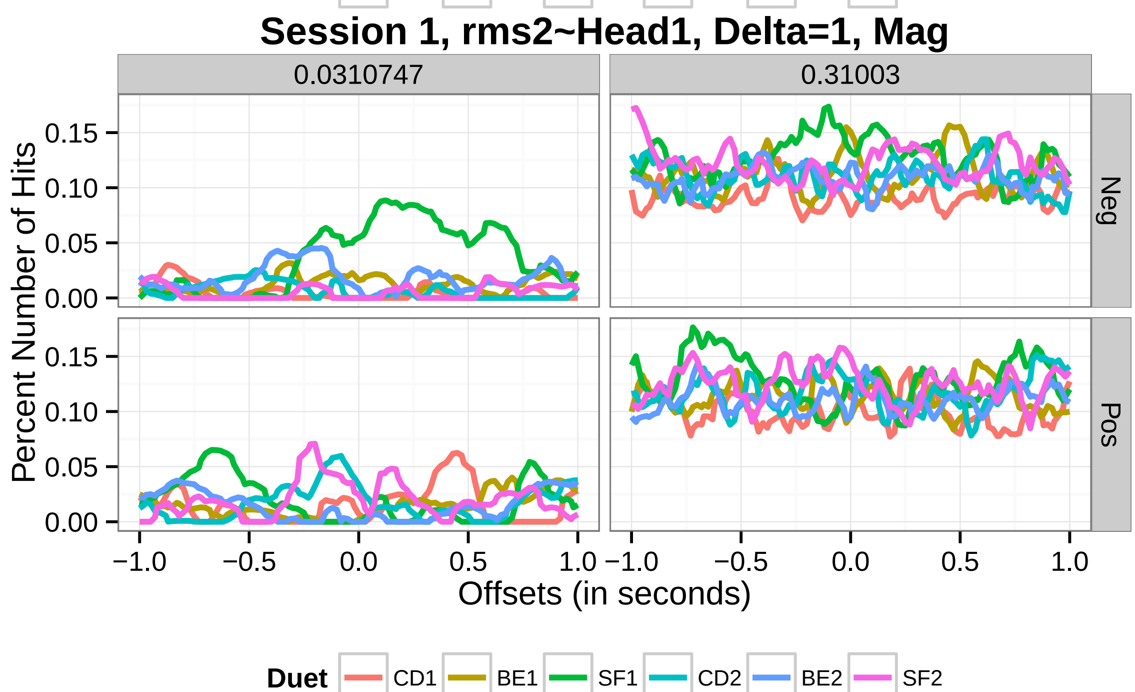


Figure B.5: Session 1 between-singer rms amplitude/head motion signal pairs

B.1.5 RMS Amplitude Pairs

The rms amplitude pairs for both Session 1 and Session 3 showed peaks in positive correlation and small troughs in negative correlation around 0 seconds offset (see Figure B.6).

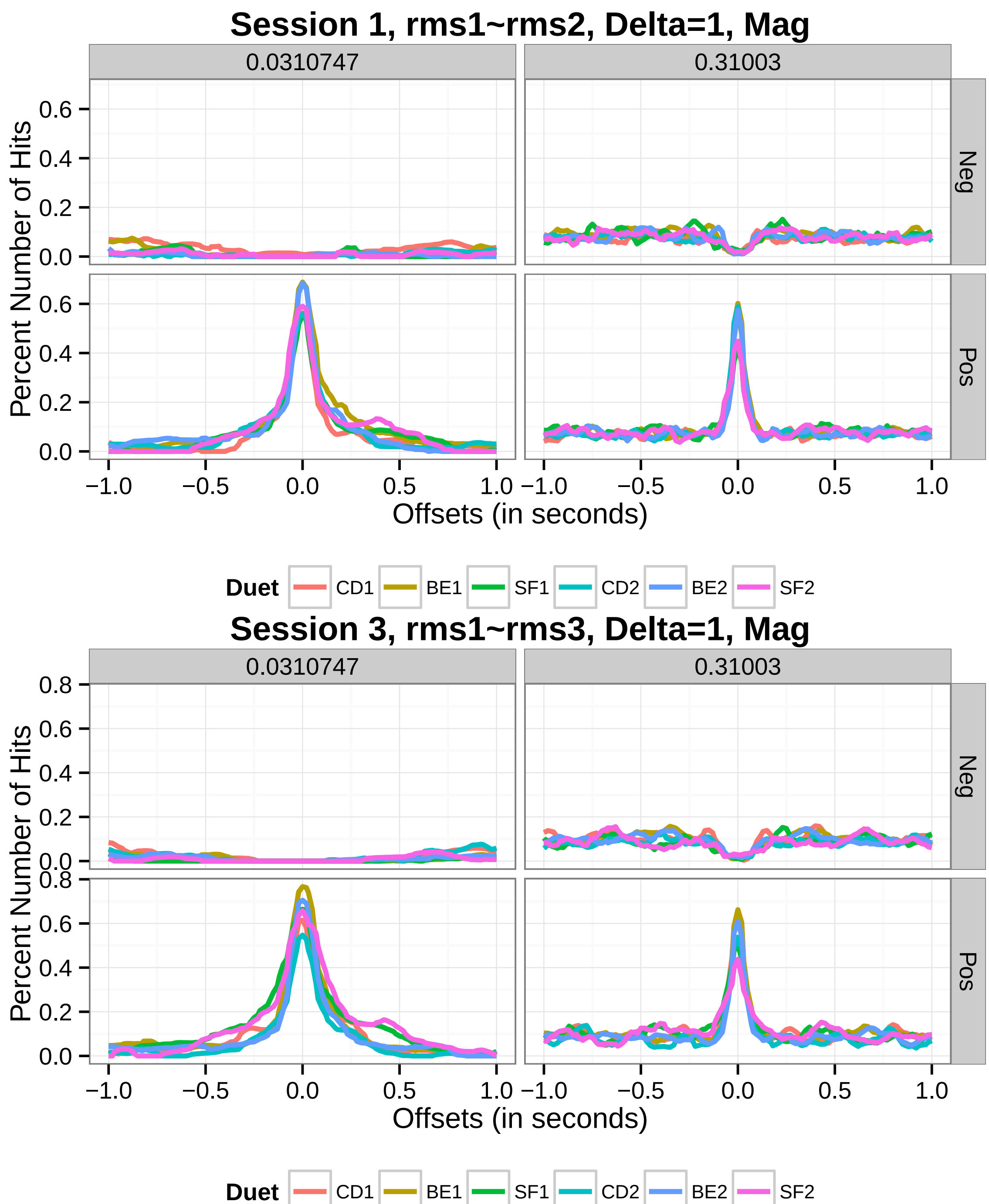


Figure B.6: RMS Amplitude Pairs

B.2 Unison Correlations

Because the whole-song correlations included both solo and unison sections, I isolated the correlation hits for the unison sections.

B.2.1 Unison Between-Singer Motion Signal Pairs

Figure B.7 shows results for Body∼Body pairs in Session 1 and Session 3. Both sessions showed high BodyA∼BodyB correlation, but not generally in the later duets.

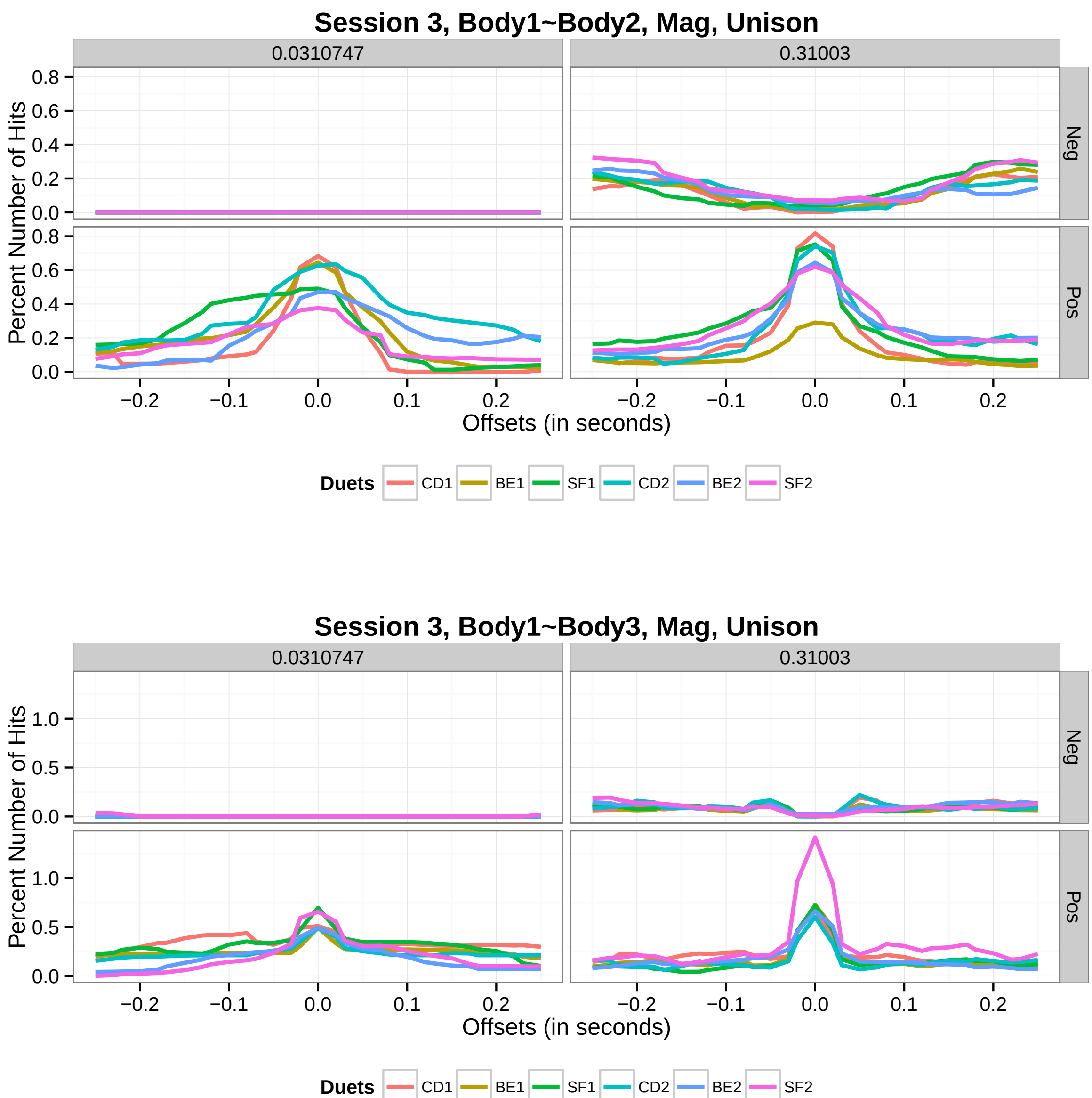
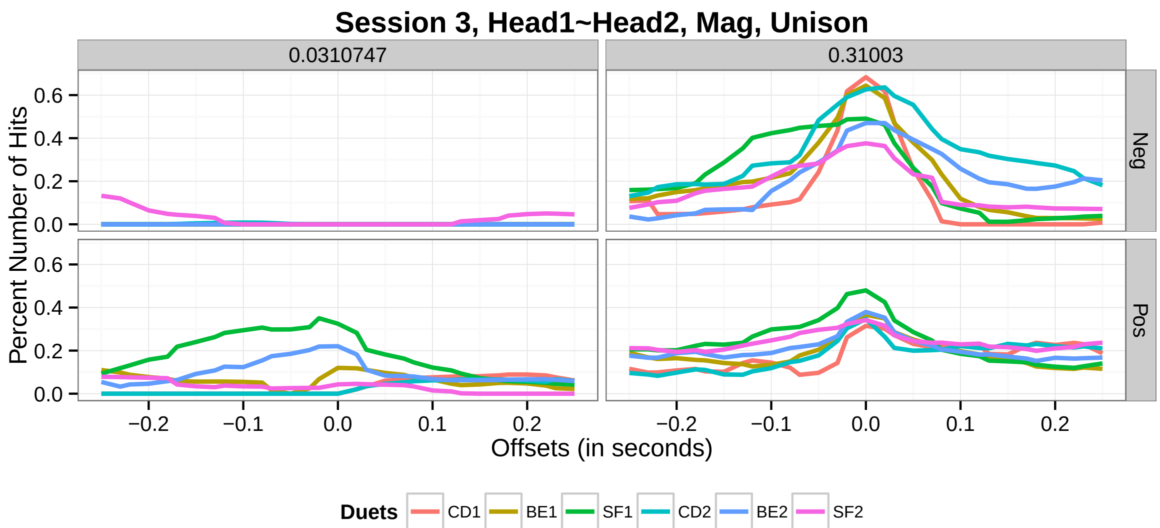


Figure B.7: Between-singer body motion signal pairs for Session 1 and Session 3

Figure B.8 shows results for Head∼Head pairs in Session 1 and Session 3. Both sessions showed peaks in negative and positive correlation around 0 seconds offset for η=0.31003. The last three duets in the session generally showed higher peaks than the earlier duets, but this didn't hold across all offsets. There were also smaller, looser peaks of positive correlation for η=0.0310747.



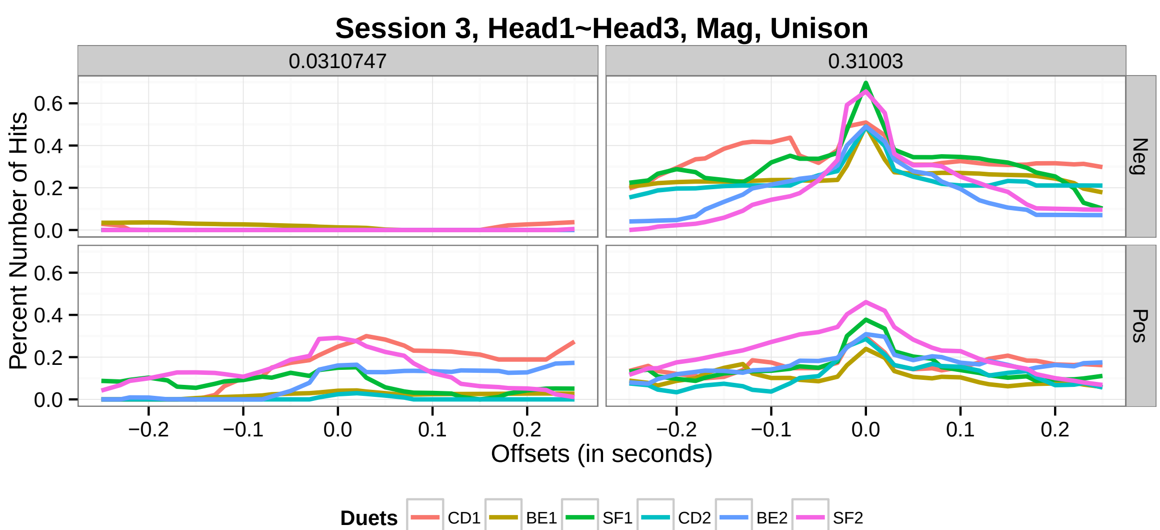
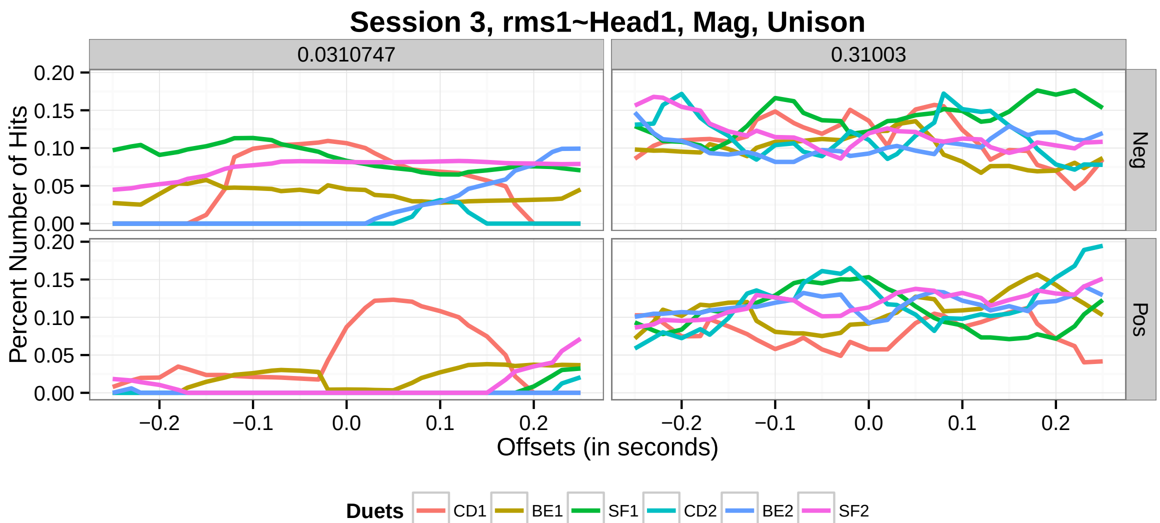


Figure B.8: Between-singer head motion signal pairs for Session 1 and Session 3

B.2.2 Within-Singer RMS Amplitude-Motion Signal Pairs

Figure B.9 shows within-singer rms amplitude-head motion pairs for Session 1. In general, correlation levels were variable and low (less than 20%).



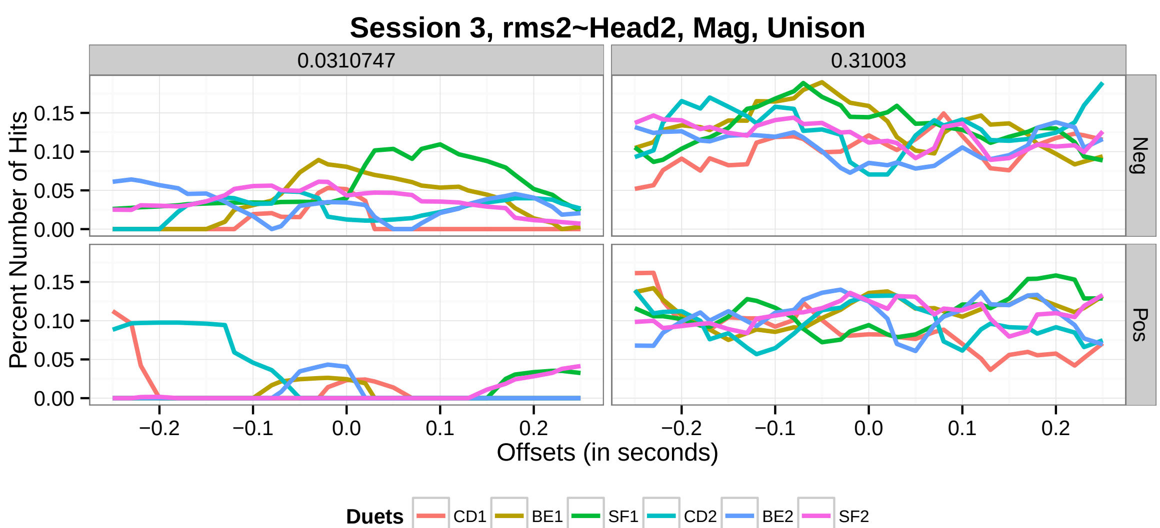
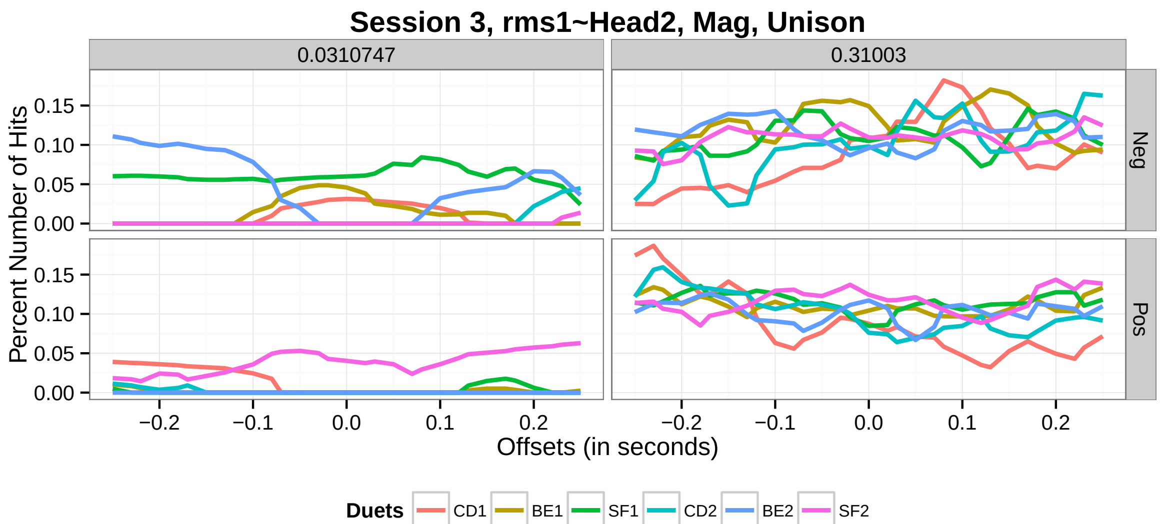


Figure B.9: Session 1 within-singer rms amplitude/head motion signal pairs

B.2.3 Between-Singer RMS Amplitude-Motion Signal Pairs

Figure B.10 shows between-singer rms amplitude-head motion pairs for Session 1. As with the within-singer pairs, correlation levels were variable and low (below 20%).



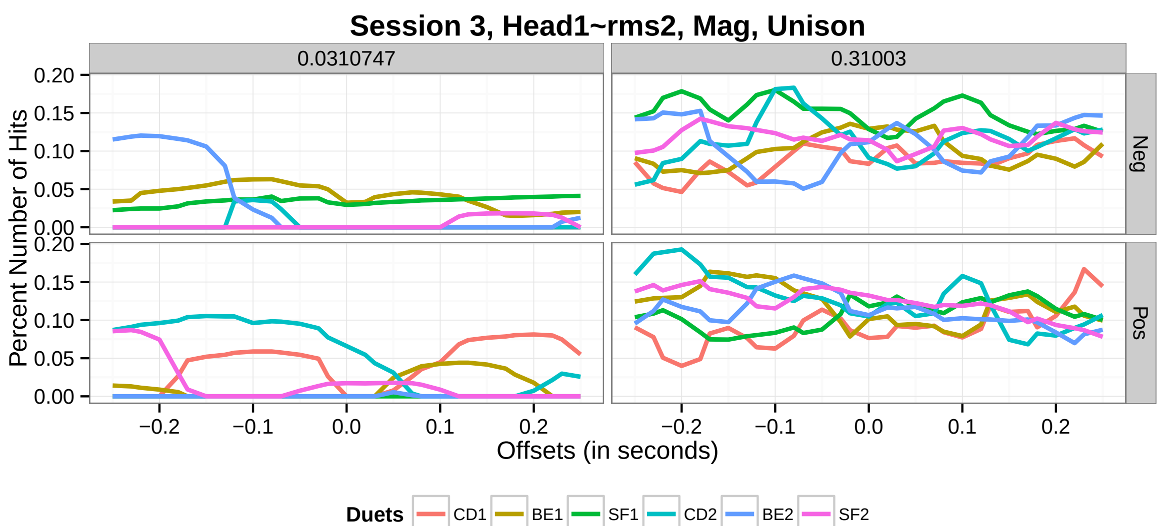
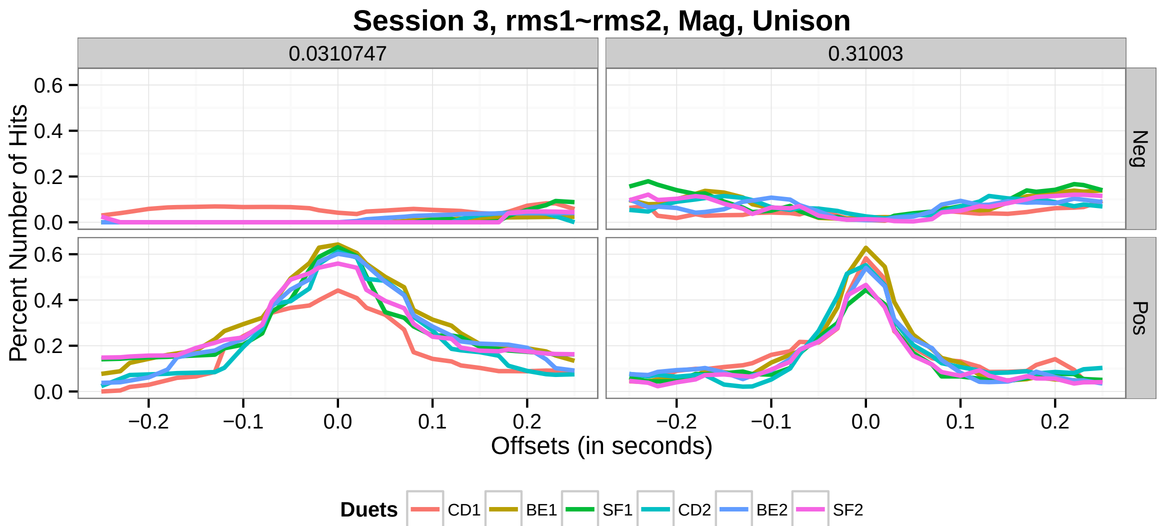


Figure B.10: Session 1 between-singer rms amplitude/head motion signal pairs

B.2.4 RMS Amplitude Pairs

Figure B.11 shows the rms amplitude pairs for Session 1 and Session 3. Overall, both pairs showed peaks in positive correlation around zero seconds offset. Later duets did not show higher levels of correlation than earlier duets, except for rms1∼rms3, which showed high positive correlations around zero offset for CD2, BE2, and SF2 when η= 0.0310747, but not when η=0.31003.



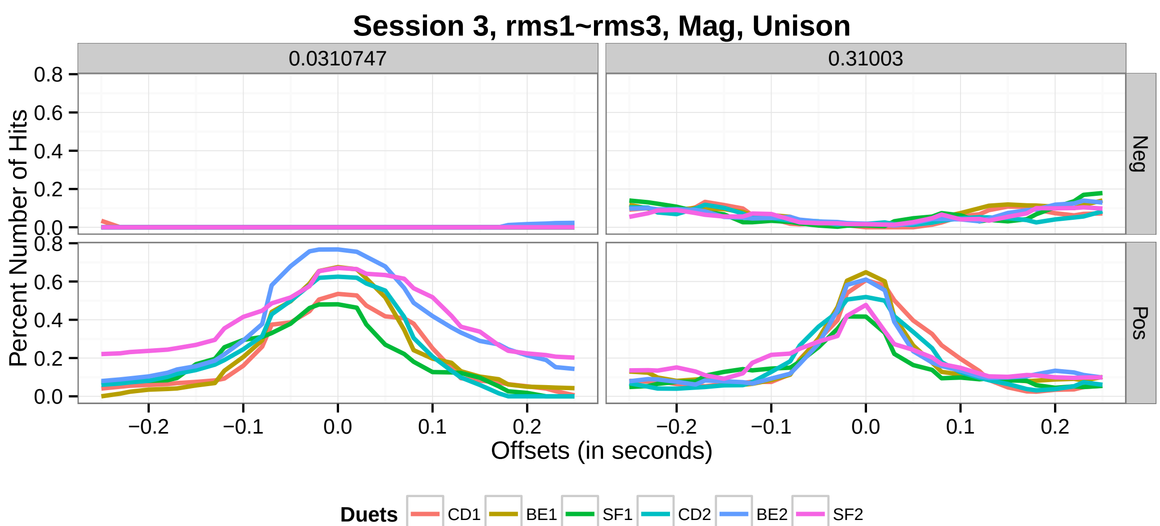


Figure B.11: Between-singer rms amplitude signal pairs for Session 1 and Session 3

B.3 Larger Offsets

I also examined larger offsets, to determine whether there were repeating peaks outside of the original 1-second offset window. Figure B.12 shows the results for the rms-amplitude pair in Session 1, with a 15-second offset window. There were large peaks in positive correlation at zero seconds offset, as seen previously, but no other large peaks beyond that.

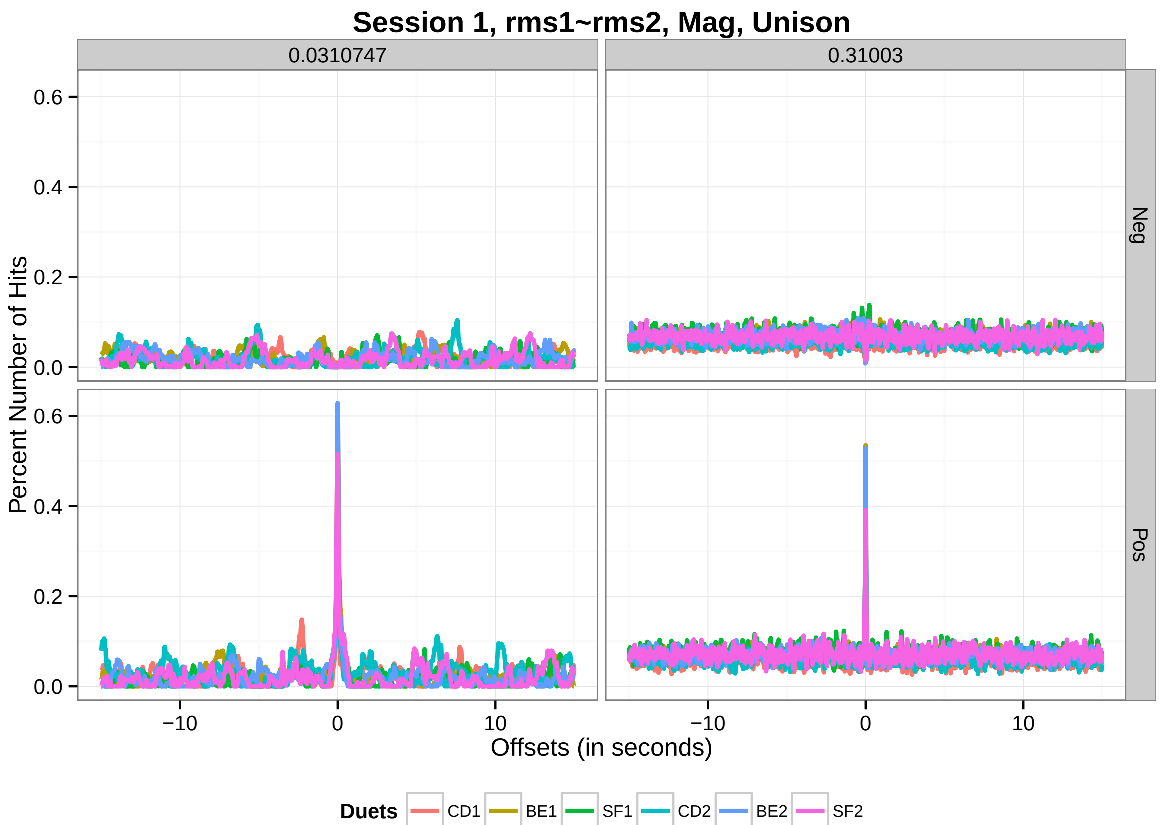


Figure B.12: Percent number correlations for rms1-rms2, 15-second offset

Appendix C

Transfer Entropy Analysis Results

This appendix contains figures (referred to in Section 3.3) showing the trajectory of ratings across duets for each Session and Token.

C.1 Unison Segments

Figure C.1 shows transfer entropy ratings for both singers across all duets and signal pairs, in both directions. There were no significant increases in transfer entropy for later duets, for any signal pair.

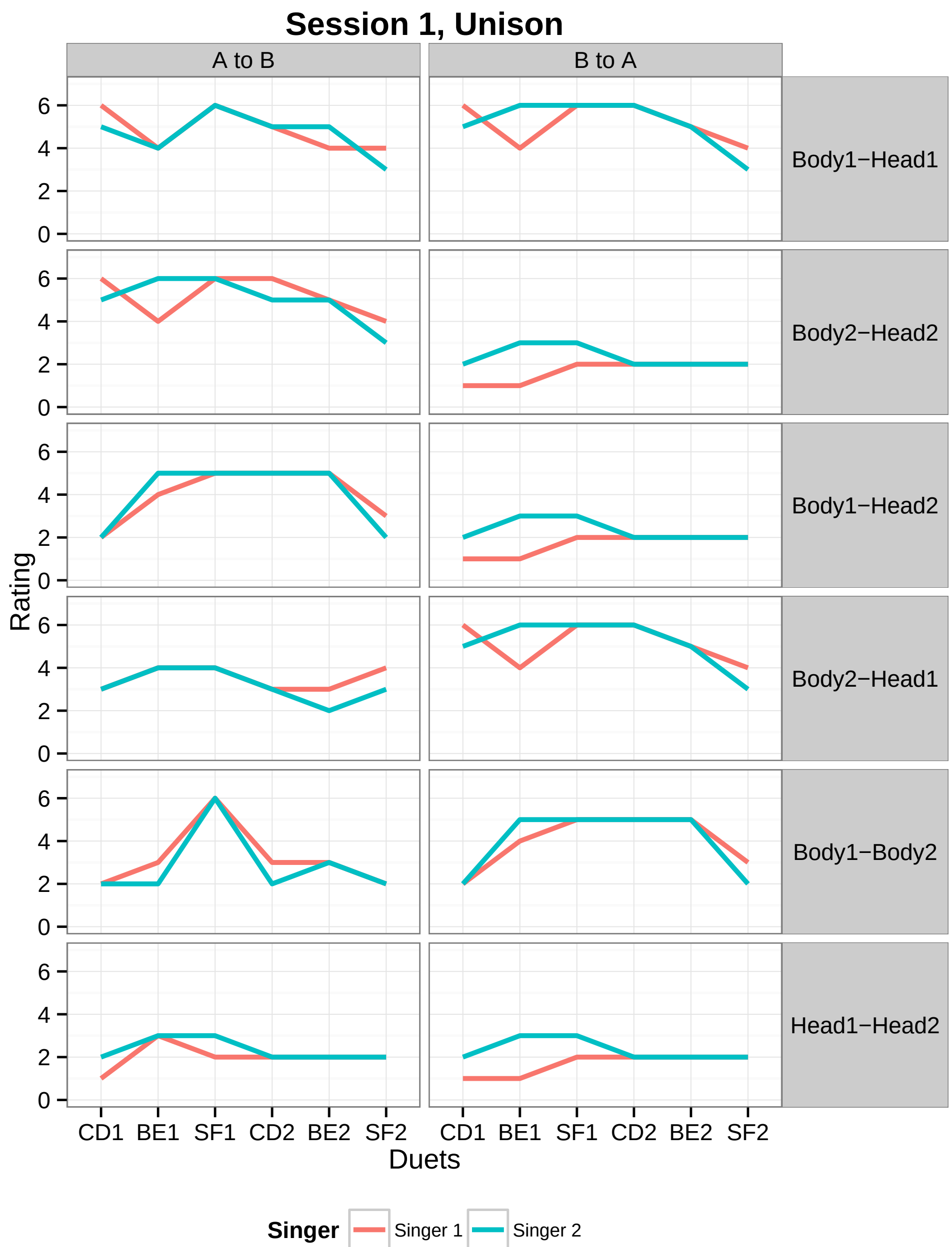


Figure C.1: Summary of Session 1 Unison Transfer Entropy Analysis Results

C.2 Solo Segments

Solo segments in Session 1 (see Figure C.2) patterned similarly to Unison segments in Session 1, in that levels of transfer entropy at the end of the session were not higher, generally, than levels of transfer entropy at the beginning. Most signal pairs showed a high level of transfer entropy for CD1 and BE1, followed by a trough around SF1 and CD1 and a small increase following BE2.

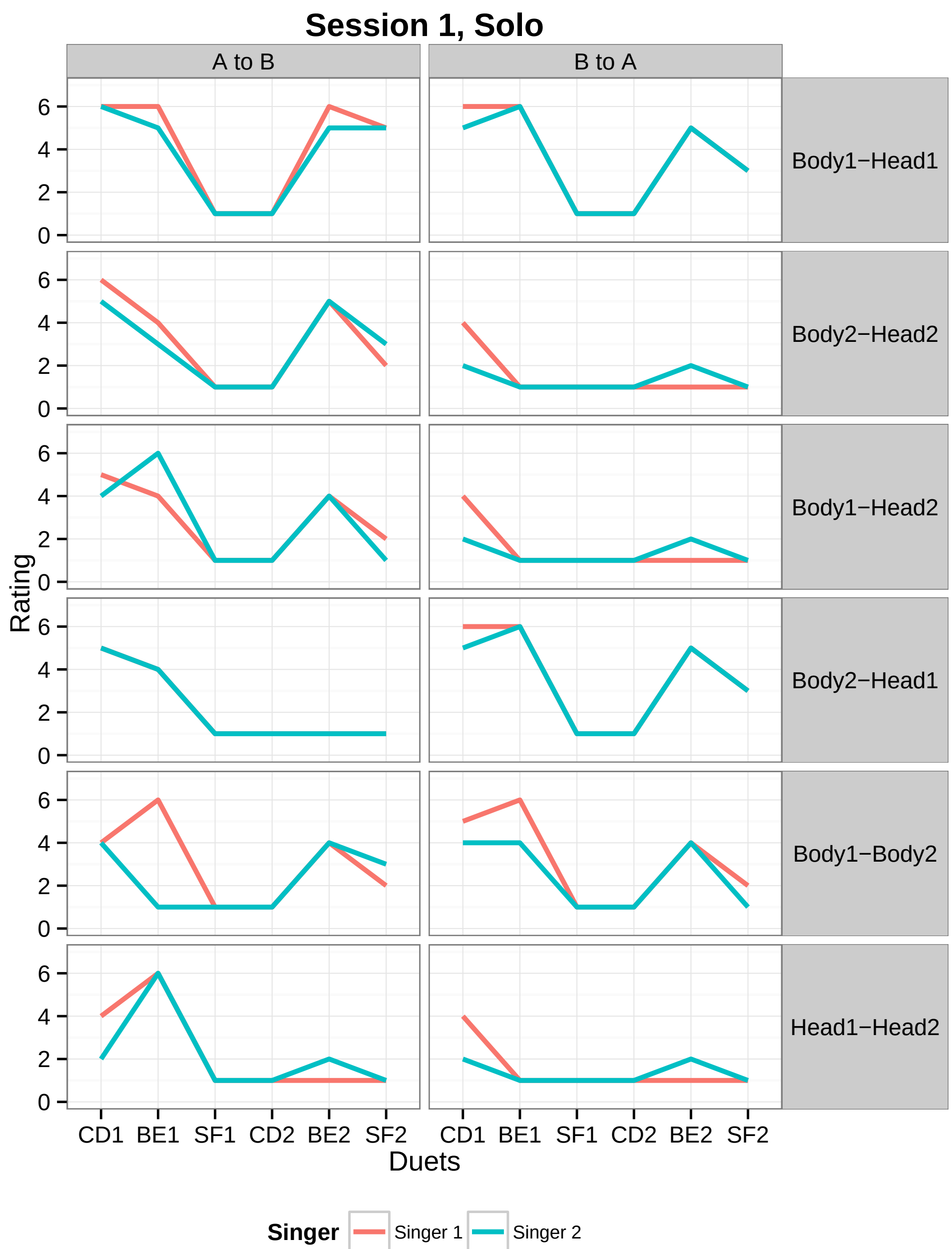


Figure C.2: Summary of Session 1 Solo Transfer Entropy Analysis Results

Session 3 (see Figure C.3 did show some increases in transfer entropy for later duets for each signal pair. The most pronounced increases were for both Body1-Head1 pairs, for Body3-to-Head3, for Head1-to-Body3, and for Body1-to-Body3.

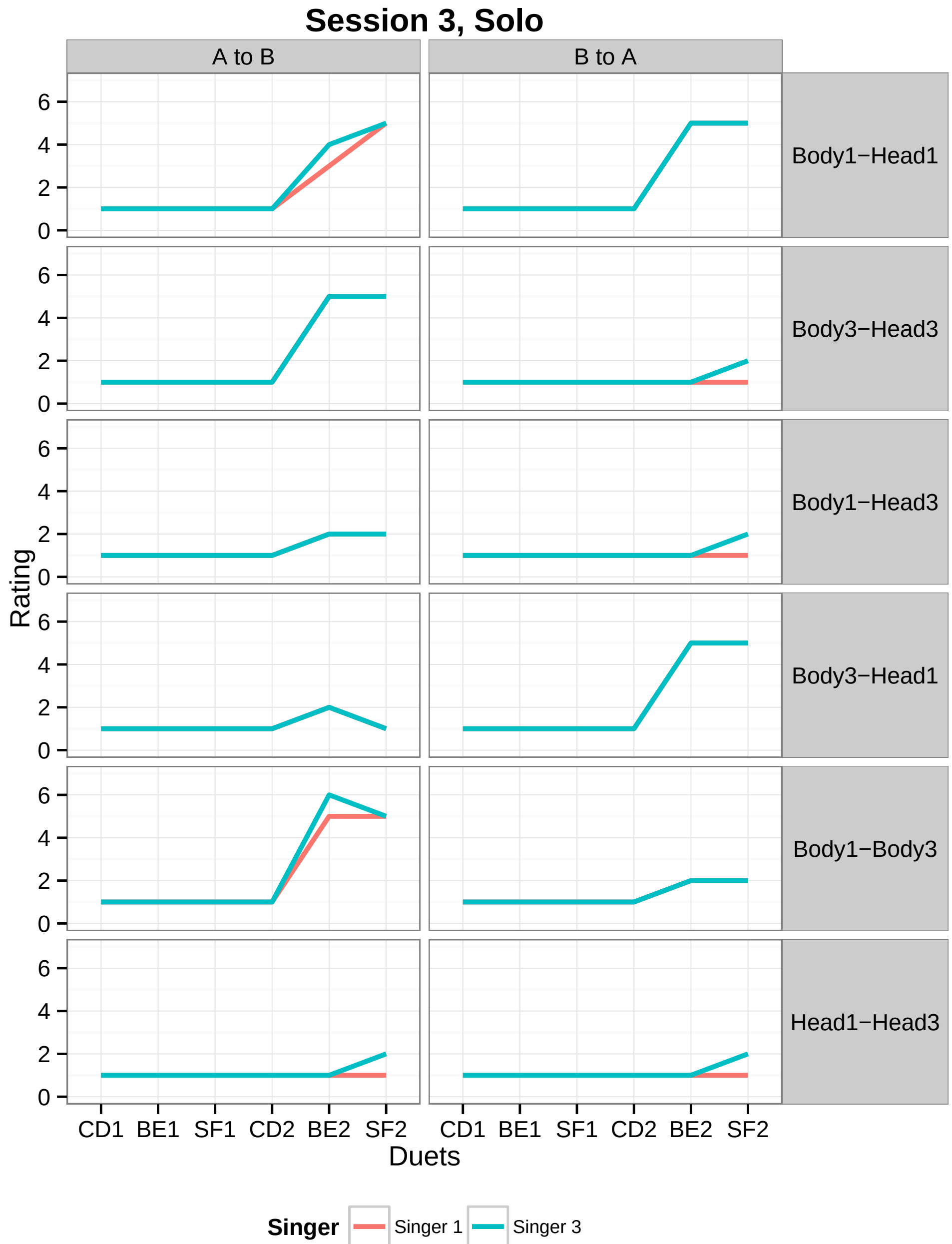


Figure C.3: Summary of Session 3 Solo Transfer Entropy Analysis Results

C.3 Unison Breaths

Session 1 again showed no consistent increases in transfer entropy in later duets (see Figure C.4). In general, SF1, CD2, and BE2 showed the highest levels of transfer entropy compared to their surrounding duets. There were, however, greater differences between specific signal pairs and between singers in for this token type than for Unison and Solo tokens. Head2-to-Body1 and both Head1-Head2 signal pairs did not show ratings higher than 3.

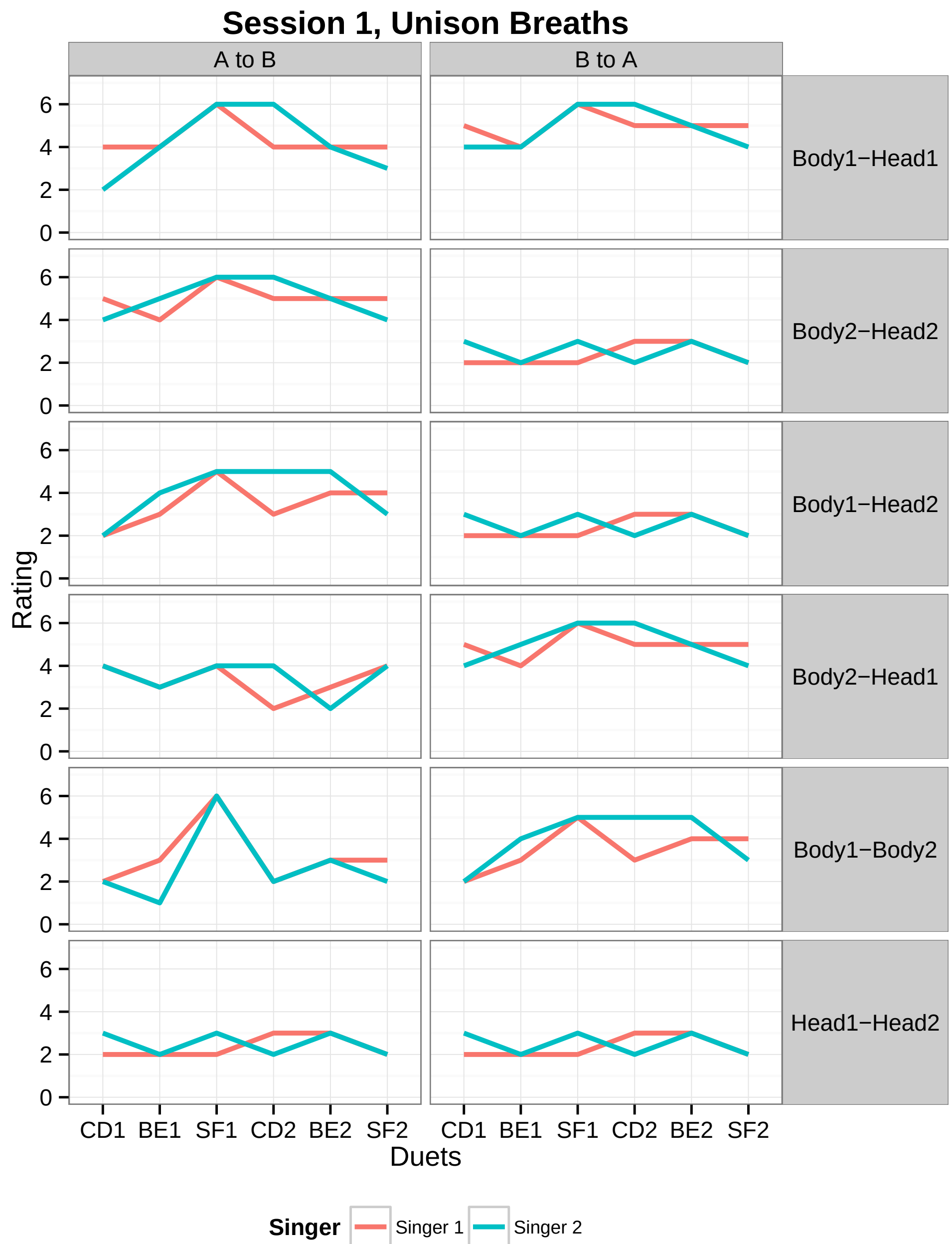


Figure C.4: Summary of Session 1 Unison Breaths Transfer Entropy Analysis Results

In Session 3 (see Figure C.5), ratings for the two singers' tokens patterned together more closely than in Session 1, with the exception of a peak for Singer 3 in BE1 for Body3-to-Head1. Singer 3 also showed increasing levels of transfer entropy in later duets for Head3-to-Body3, both Body1-Head3 pairs, Body3-to-Body1, and both Head1-Head3 pairs. Singer 1's results were more variable across duets.

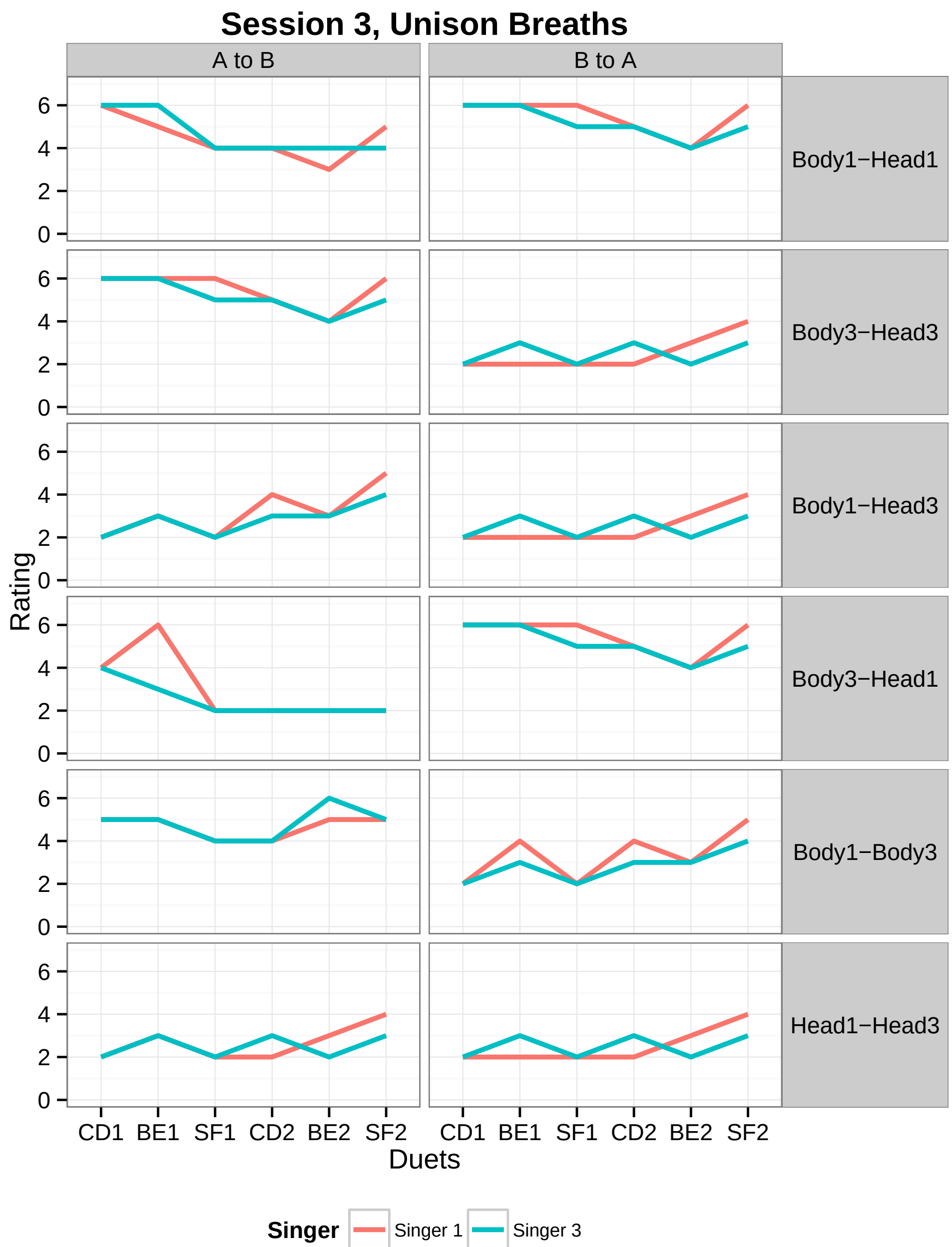


Figure C.5: Summary of Session 3 Unison Breaths Transfer Entropy Analysis Results

C.4 Solo Breaths

Similarly to breaths before unison segments, breaths before solo segments in Session 1 (Figure C.6) showed higher levels of transfer entropy for earlier and later duets, with lowest levels of transfer entropy around SF1 and CD2. Peak levels of transfer entropy rated around 6. Head2-to-Body2, Head2-to-Body1, and both Head1-Head2 signal pairs showed the lowest levels of transfer entropy across all duets, consistently rating no higher than 2.

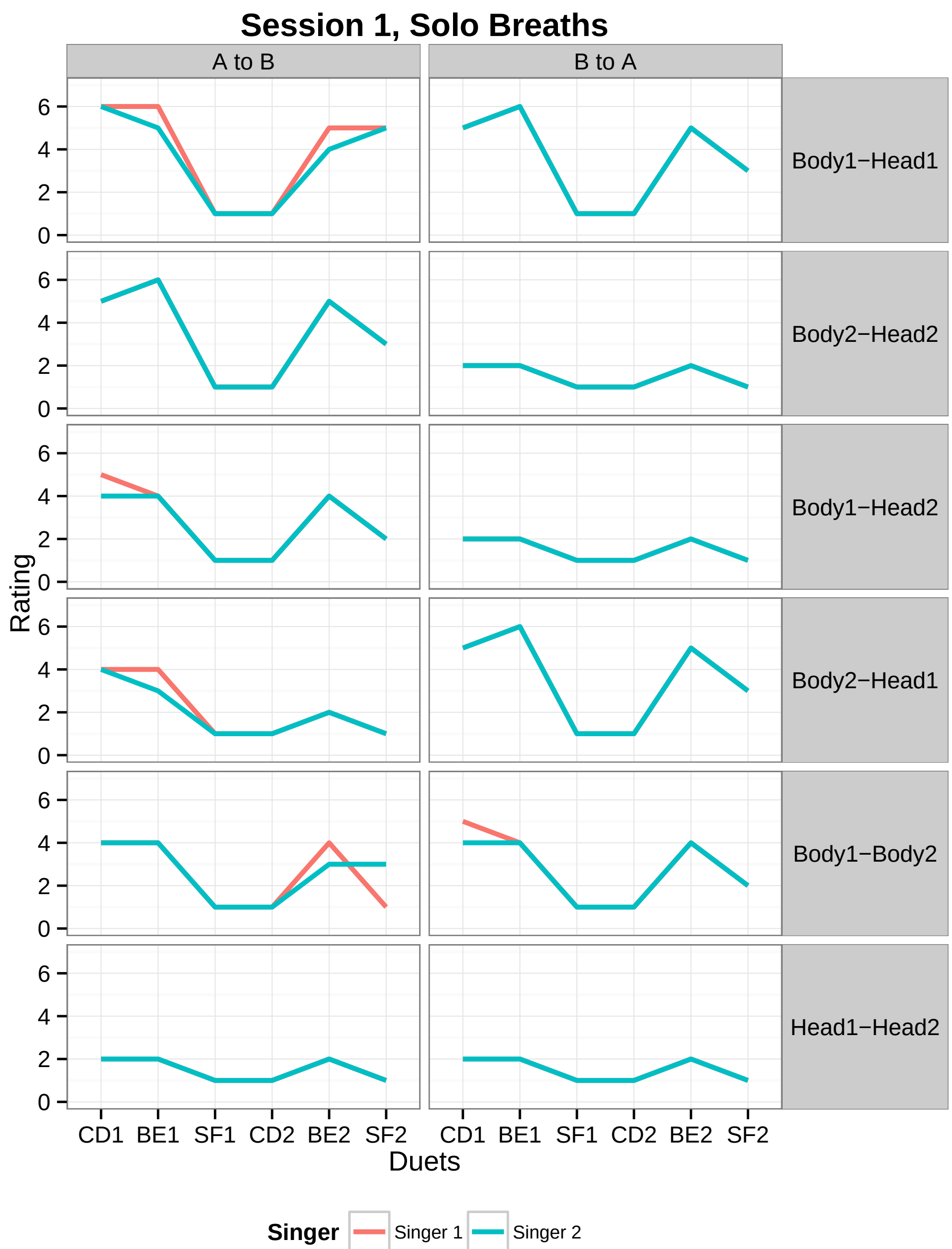


Figure C.6: Summary of Session 1 Solo Breaths Transfer Entropy Analysis Results

In Session 3 (Figure C.7), later duets did show higher levels of transfer entropy than earlier duets. The earliest duets consistently rated 1. Transfer entropy levels increased earlier for Singer 1 than for Singer 3 (at CD2). Singer 3 did not show any change in level of transfer entropy across all duets for Head3-to-Body3, Head3-to-Body1, Body3-to-Head1, and both Head1-Head3 signal pairs.

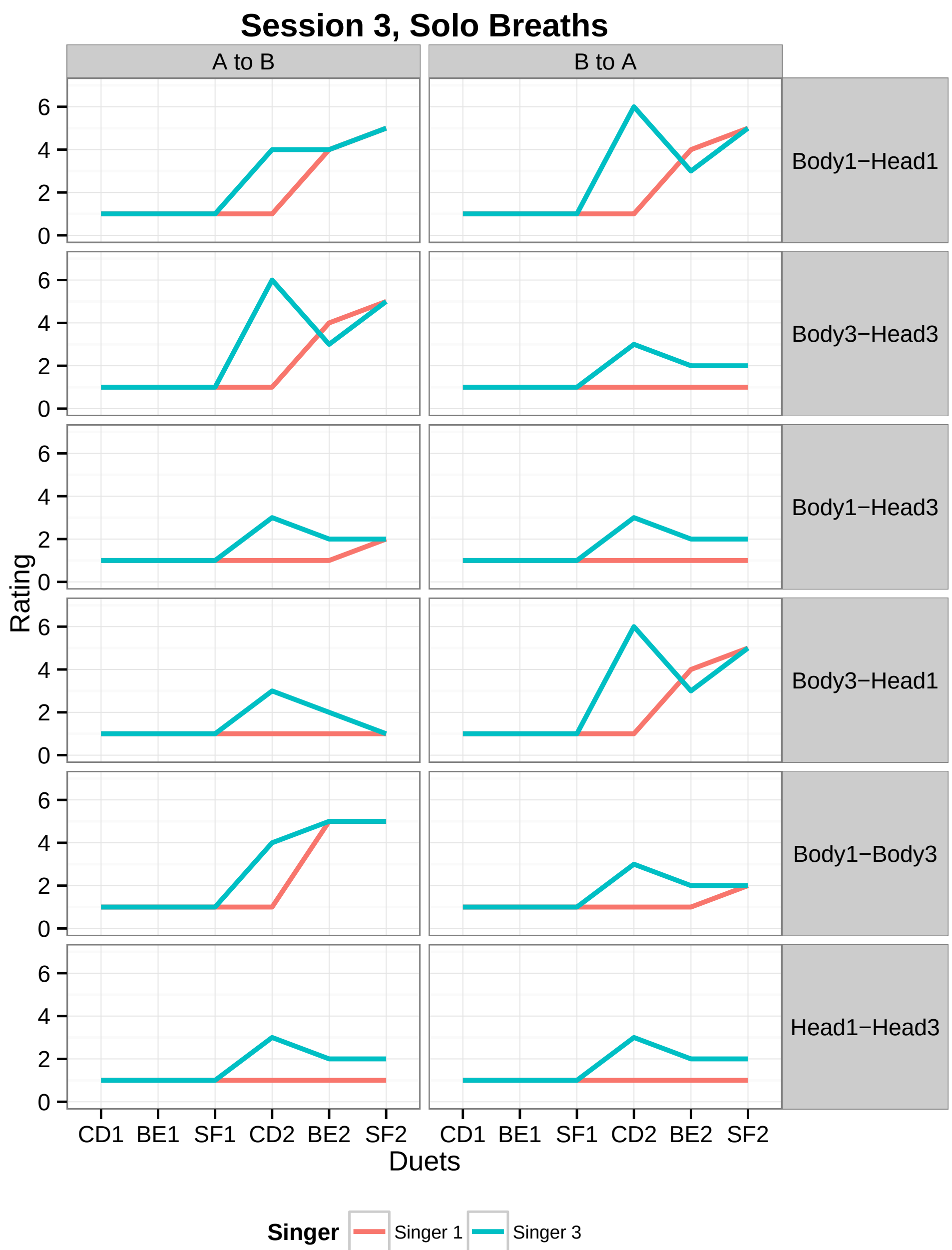


Figure C.7: Summary of Session 3 Solo Breaths Transfer Entropy Analysis Results

C.5 Looks At

In Session 1 (Figure C.8), Singer 2 showed somewhat consistent increases in levels of transfer entropy in later duets. Singer 1 showed identical patterns in levels of transfer entropy across all signal pairs–a peak of 3 for BE1 and SF1, followed by a drop off at CD2, where there were no Looks At tokens.

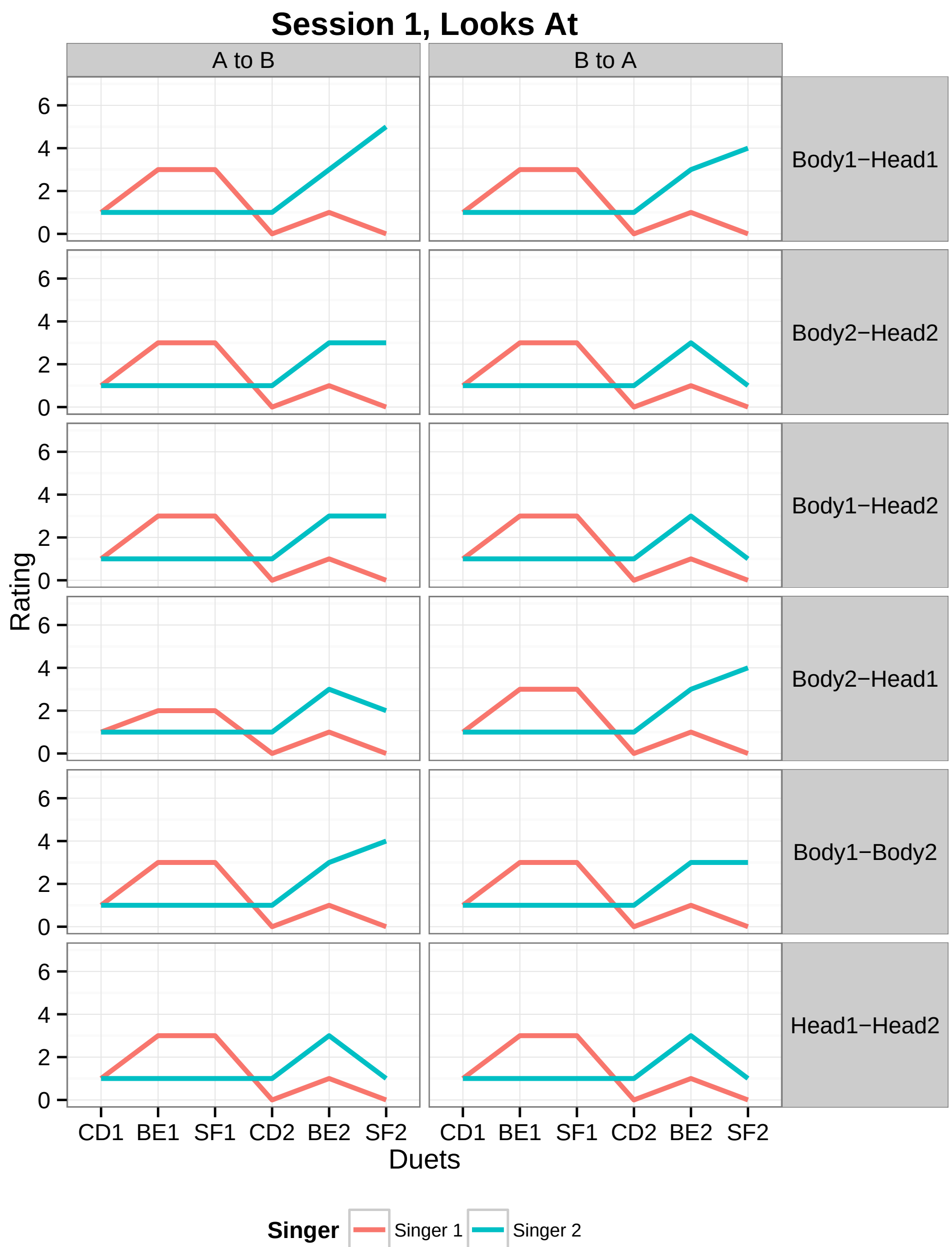


Figure C.8: Summary of Session 1 Looks At Transfer Entropy Analysis Results

In Session 3 (Figure C.9), all results were more or less identical; low levels of transfer entropy (rating at 1) for CD1 and BE1 for Singer 3, and for CD1, BE1, and SF1 for Singer 1. There were no tokens recorded for the remaining duets.

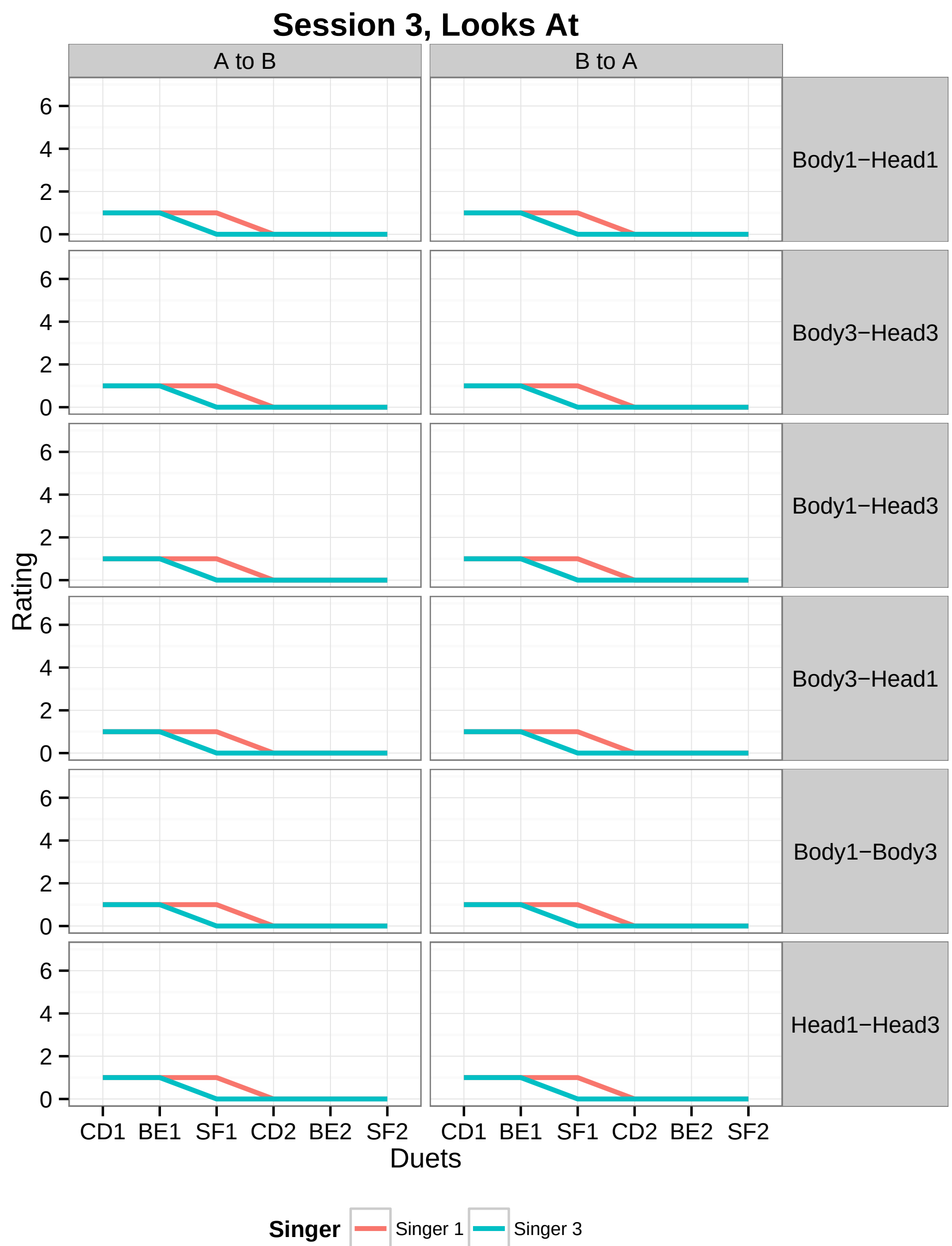


Figure C.9: Summary of Session 3 Looks At Transfer Entropy Analysis Results

C.6 Looks Up

In Session 1 (Figure C.10), there was in general a peak for Singer 1 earlier in the session (at BE1) and a peak for Singer 2 later in the session (at BE2). These peaks were highest for both Body1-Head1 pairs, Body2-to-Head2, and Head1-to-Body2. Singer 1's levels of transfer entropy did not peak at all for Head2-to-Body2, Head2-toBody1, or for either Head1-Head2 pair.

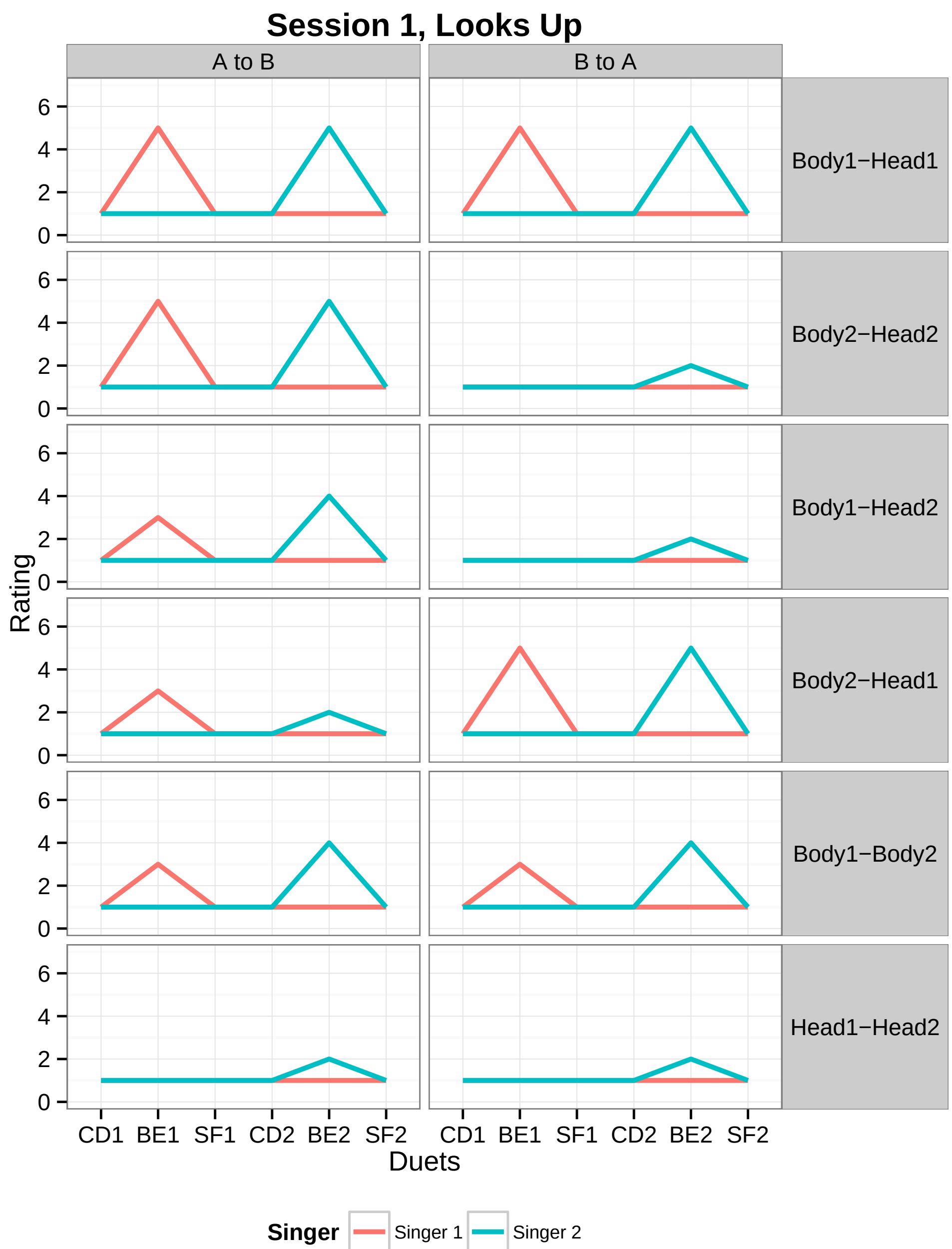


Figure C.10: Summary of Session 1 Looks Up Transfer Entropy Analysis Results

Session 3 (Figure C.11) showed generally higher levels of transfer entropy than Session 1. Singer 1 tended to peak at a rating of around 5 or 6 in the first three duets, sometimes with a secondary peak before CD2. There were no tokens recorded for Singer 3 in duets CD2, BE2, and SF2. Singer 3 tended to show two peaks in transfer entropy levels, one at BE1 and one at BE2. These peaks were highest for both Body1-Head1 signal pairs, for Body3-to-Head3, for Head1-to-Body3, and for Body1-to-Body3.

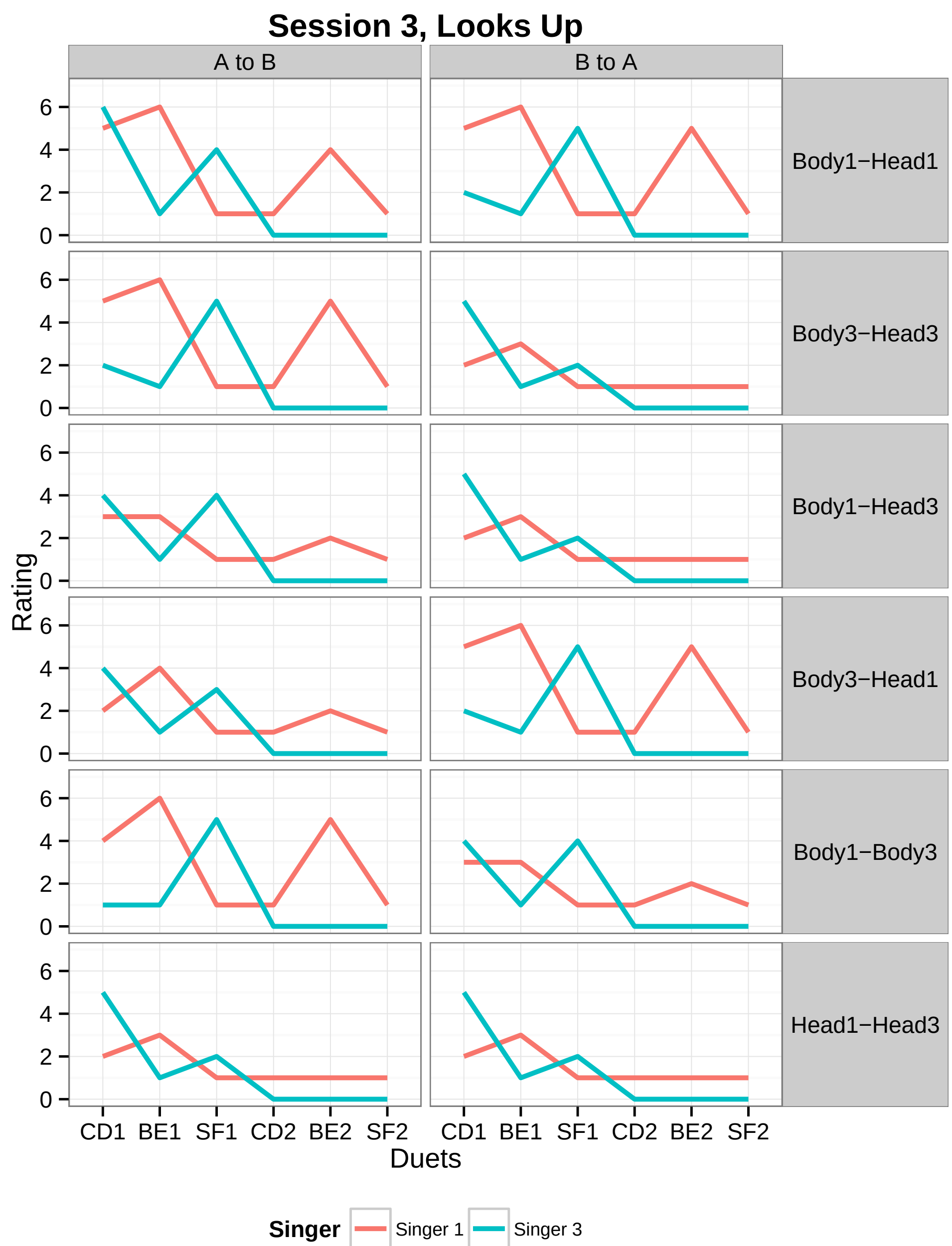


Figure C.11: Summary of Session 3 Looks Up Transfer Entropy Analysis Results