

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Using Music as a Turn in Conversation in a Lesson

Permalink

<https://escholarship.org/uc/item/5n40g8kk>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN

1069-7977

Authors

Duffy, Sam

Healey, Patrick

Publication Date

2013

Peer reviewed

Using Music as a Turn in Conversation in a Lesson

Sam Duffy (s.duffy@eecs.qmul.ac.uk)

Patrick G. T. Healey (ph@eecs.qmul.ac.uk)

Cognitive Science Research Group

School of Electronic Engineering and Computer Science

Queen Mary University of London

London, E1 4NS, United Kingdom

Abstract

Music is sometimes compared to language as a system of communication, however this comparison is usually at a generic formal, cultural or social level. This paper explores this analogy at the detailed level of interaction: to what extent can musical contributions act as conversational turns? We explore this question through an ethnographic study of music lessons. We describe a new transcription notation designed to capture the interactional details of musical contributions. Using this notation we show that although the ultimate objective of a lesson is development of musical performance, the detailed structure of the musical contributions depends on their interactional organisation. We show that musical contributions display interactional structure at the turn and sub-turn level and are closely integrated with other verbal and non-verbal cues as part of the unfolding conversation.

Keywords: music tuition; conversational turn; interaction; repair

Introduction

Comparisons of music and speech are well documented, for example Feld and Fox (1994) critically review a broad interdisciplinary collection of anthropological work on the relationship of music to language. Besson, Chobert, and Marie (2011) take a cognitive approach, considering evidence for the bidirectional influence of musical expertise on speech processing and of linguistic expertise on the processing of harmonic sounds. Zatorre, Belin, and Penhune (2002) show that whilst each auditory cortex, in the left and right hemispheres of the brain, has been shown to favour processing of either music or speech, this complimentary specialisation should “*be seen as arising from a single underlying principle rather than being unrelated phenomena*”.

In this paper we look at the relationship between music and speech in social interaction through the study of instrumental music lessons. We describe a new transcription notation designed to capture the interactional details of musical contributions. This is used to investigate the extent to which musical utterances produced during a lesson act as conversational turns. Initially we will summarise some of the characteristics and rules which govern turn taking in conversation. We will then look at the different, context-dependant roles that music can play in different types of interaction. We will use an ethnographic study of clarinet lessons to provide examples of the interplay between music and speech in a pedagogical setting

What is a Turn?

In conversation, the participants manage their exchange of units of speech. Sacks, Schegloff, and Jefferson (1974) set out

rules describing the mechanics of how this is achieved, identifying dialogue as a turn organised activity. In order to investigate any turn taking system, be it playing a game of cards or managing a queue, there is a need to define what constitutes a turn. In conversation analysis, the building block for turns is the turn construction unit (TCU), which can be formed from a single word or utterance, a sentence, or a phrase. As one speaker approaches the possible completion of a TCU, another speaker may recognise this as a transition-relevance place (TRP) where they can take the floor. However this is an opportunity rather than an obligation. It may be that the current speaker starts a new TCU and continues with their turn. Since it is possible to predict when turns are heading towards completion, the next speaker can often start their turn without a perceivable gap in the conversation, or even start before the current speaker has finished, causing a brief overlap. Whilst we do not usually talk at the same time as someone else for prolonged periods, brief overlaps like this at transition points are frequent. To investigate these rules in action, a set of notations was proposed for use in conversation analysis transcripts (for example see Figure 1).

```
// The point in the current speaker's speech where overlap
   by the other starts
* The point where overlap finishes
= latching (no interval between two pieces of talk)
() elapsed time in tenths of seconds, used to denote
   pauses and silence
: Prior syllable is prolonged
( ) empty parenthesis indicate an utterance which could not
   be clearly heard
(( )) double parenthesis indicate features other than
   verbalisation
```

Figure 1: Conversation analysis notation (Sacks et al., 1974)

Non-verbal interaction

The gestures which accompany speech are an important scaffold to conversation as well as an integral part of interaction. Bavelas, Chovil, Lawrie, and Wade (1992) explain that conversation is not made up of alternating monologues but is an interactive social system, and show that interactive gestures are essential in maintaining conversation. Cassell and Thorisson (1999) describe the importance of *envelope* feedback, non-verbal accompaniments to speech such as beat gestures, gaze and head turns. Engle (1998) demonstrate that when gesture and speech are consistent with respect to the underlying referent, they are understood as composite signals rather than separate channels. Clark and Krych (2004) show the im-

portance of visual modality to collaborative work with shared objects, gesturing in relation to an object whilst another is speaking effectively reducing the number of turns required for task completion. Non-verbal interactions can also function as a turn in themselves, the expected response to a verbal turn taking the form of physical activity rather than speech. For example, a request at the family dinner table during a busy multi-party conversation to pass the butter can be actioned without any reference in the continuing conversation (E. A. Schegloff, 2007, pp.10 *Chicken Dinner*).

In the same way that gestures act as a scaffold to speech, non-verbal interaction is vital to the co-ordination of musical sound production. In jazz, a solo may be scored as for a prescribed number of bars with a chord structure, but it may also be open, thrown around the group for others to take a turn or even choose to play against each other, so non-verbal communication with fellow performers is important to manage improvisation. Observing a group of musicians engaged in free improvisation sessions, Healey, Leach, and Bryan-Kinns (2005) found that the musicians used the patterns of body position and orientation, or f-formations (Kendon, 1990) typical of face-to-face conversation to organise the timing of their musical contributions. Moran (2011) observes a group of North Indian musicians whose vocal and bodily responses to musical ideas could be interpreted as comparable to the function of back channelling in everyday conversation. Even in the performance of a predetermined composed score, an ensemble must synchronise entrances and exits, and changes in dynamics and tempo and this is usually achieved by gestures, head and body movement and gaze. In an analysis of co-ordination between members of a string quartet, Davidson and Good (2002) wrote "*The nature of the interaction and co-ordination in conversation are, we believe, analogous to that in small group music-making contexts.*"

Turn breakdown and self repair

An important part of maintaining conversation is dealing with turn-taking errors or rule violations such as misunderstandings, interruptions, gaps and overlaps. These are frequent in natural dialogue and we manage these breakdowns through repair. Self-repair occurs when the current speaker manages an error within the same turn in which it was made. To do this, the speaker must be able to self-monitor and detect a problem in what they are saying, or see some outward sign of the listener's confusion. Correction is made promptly once the problem has been detected, a neutral holding term such as 'uh' often being used to communicate error detection to the listener and so hold the turn for the self-repair to be made (Levelt, 1983). Self-repair occurs much more frequently than other-initiated repair in natural speech. However even if the other speaker initiates repair, they are much more likely to encourage the original speaker to correct themselves, rather than make a direct correction (E. Schegloff, Jefferson, & Sacks, 1977).

The Role of Music in a Pedagogical Context

There are many different types of interaction involving the production of music. From an ethnomethodological perspective, the social norms which govern how members of a group understand their world and so behave in it (Garfinkel, 1964) are relevant to any interaction, whatever the mode of communication. The context of a musical interaction is therefore important in determining the roles of the participants and the communication content of the music produced. Players in an ensemble rehearsal will use verbal interaction to analyse, discuss and shape their approach to a piece. During a performance, they cannot use the same level of verbal interaction unless they want to share their inner workings with the audience. A soloist is not generally expecting to enter into a dialogue with the audience, however they may engage in extensive non-verbal interaction with fellow performers and the conductor, as part of co-ordinating their performance with an orchestra.

In a pedagogical context, such as an instrumental lesson, both student and tutor produce music but it is subject to immediate scrutiny, their musical utterances being produced with the expectation of immediate feedback. The tutor is not listening to the performance from the perspective of an audience member, but as an expert critic and must be able to immediately verbalise their assessment of the student's performance. In order to prepare a complete piece or movement, they will focus on a small part of it each week, building up the work gradually. Small fragments of music, perhaps only a few bars, are worked on at a time. The student plays them, receives feedback, then plays them again for the tutor to assess if they have incorporated the feedback, and so on in an iterative process.

Representing Music in a transcript

It is surprising how few authors looking at interaction in a musical context attempt to represent the musical sounds produced. When they do, one approach is to use musical notation to locate activity on a musical timeline, for example Figure 2 (Holck, 2002). However this presentation is less meaningful for those who do not read music.

In applying the rules of conversation analysis to musical utterances, our notation needs to be comprehensible to those who are used to working with transcripts, whilst being able to capture the interactionally relevant aspects of music production. A system has been devised with two main aims: to make representation of the music understandable for both musicians and non-musicians, and make it possible for a written transcript to convey the full interaction whether utterances are verbal or musical. The starting point was established notation for conversational analysis as shown in Figure 1. This was adapted to produce notation for musical sounds as shown in Figure 3.

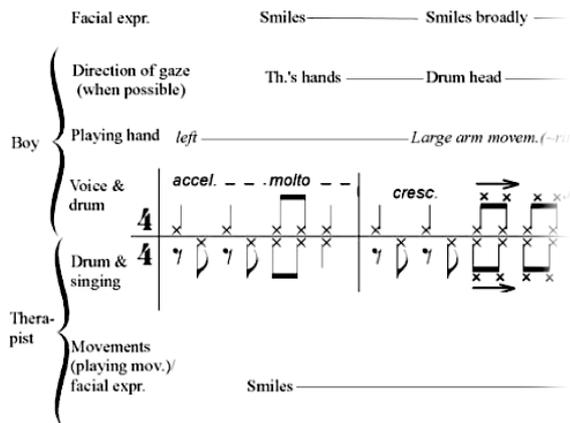


Figure 2: Extract from a transcript using musical notation. English translation used with permission Holck (2007, pp 33)

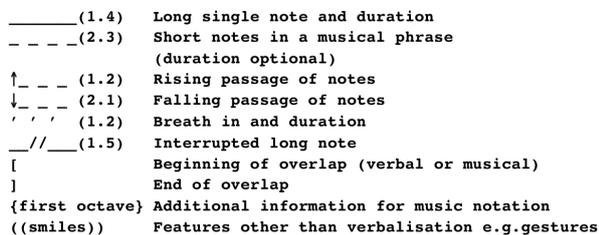


Figure 3: Notation devised to represent musical sounds.

Applying Musical Notation to Data

We have observed clarinet lessons with four students and two tutors at two music schools in London which prepare young adults for the study of performance at an undergraduate level. For explanatory purposes we focus on the turn-like organisation of musical contributions in one lesson however, we have found that these phenomena recur across the sample. The student is a male clarinet player studying for his grade eight exam. Throughout, we will indicate the areas of interest which are to be the subject of further work using the broader dataset which has been collected.

Non-verbal interaction as a scaffold for musical utterances

The student is being tested on playing scales from memory. The tutor holds a small book containing the syllabus in her hands, at chest height. The student is holding his clarinet in front of his body with both hands and they are facing each other. The tutor has asked him to play a three octave scale. There are eight notes in an octave, the last note and the first note are the same, the last note being an octave higher in pitch. To simplify our transcript, we can represent this as numbers rather than notes 1 2 3 4 5 6 7 8(1), where 8 represents the top of the scale. To play a scale over several octaves you would play seven notes, then start the next octave so a three octave scale could be represented as 1 2 3 4 5 6 7 1 2 3

4 5 6 7 1 2 3 4 5 6 7 8. The instances of octaves are shown separately on the transcript below for clarity. The student (S) is expected to play all three octaves ascending and then descending in one smooth phrase. He commences the ascending scale, flicking his eyes up to the tutor (T) occasionally, but makes an unpleasant squeak at the start of note 7 of the third octave. He elongates this note and then briefly stops playing (Figure 4).

T: OK great (0.33) Ah let's have (0.44) F sharp major ((T moves head up from book to look at student)) (0.82) tongued ((S brings clarinet up to mouth))

T: so by tongued I mean legato [tongued you know
S: [yeah

((T looks back down to her book))

S: ' ' '(0.82)

=↑1_2_3_4_[5_6]_7_ {first octave} [S glances briefly up at tutor]

=↑1_[2]_3_4_5_6_7_ {second octave} [S glances briefly up at tutor]

=↑1_2_3_4_5_6_ {third octave}

={squeak} 7 _____ (0.75)

pause (0.16)

=↑ [8 _____] (0.75) [S looks up at tutor and raises eyebrows]

S: the[re? ((speaks with clarinet still in mouth))
T: [because you want E sharp here

Figure 4: The student seeks guidance musically and verbally.

The tutor does not change her position or gaze, or seek to interrupt in this brief pause, and the student continues with the next note in the scale (note 8) again elongating it, looking up at the tutor as he plays and raising his eyebrows. This is a possible parallel with Levelt's word-completion hypothesis (Levelt, 1983) in which it is suggested that speakers have a tendency to complete words after detection of trouble. Musicians may exhibit a similar tendency to finish a musical idea, although this does not always happen, since examples of the student restarting mid-phrase after a mistake have also been found. The circumstances around whether the student finishes a phrase or restarts mid-phrase is a potential area of further investigation. The tutor still does not visibly react. Rather than continue with the scale (which would now be the three octave descent), the student stops and verbally seeks guidance 'there?' with the clarinet still in his mouth (Figure 5). Finally she responds, verbally indicating that the error was with the note that played by not referring to the squeak.

The tutor puts her book down on the music stand, turns away and picks up her clarinet from its resting place on the piano. She then twists her body back to look up and meet the student's gaze as he retries the top of the scale tentatively (Figure 6). They continue to make frequent brief direct eye contact and the tutor nods and encourages him verbally 'yeah' as he starts the descent. He continues down the scale, picking

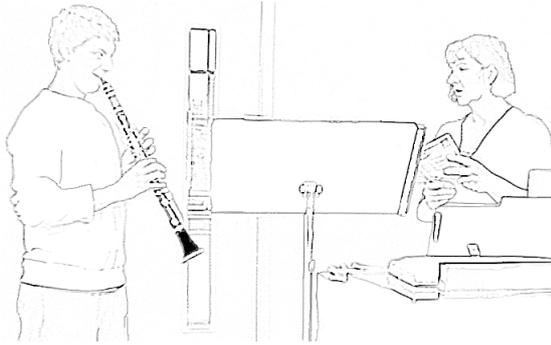


Figure 5: The student seeks verbal guidance "There?"

up speed to finish with a confident long final note. The tutor continues to hold her clarinet close to her mouth but does not play, letting him finish before giving verbal feedback on the rest of the scale.

```
S:  _ _ _ _ ^7_8_ {resumes third ascending octave}
S:  =^7_6_5_ {starts third octave descending,
tentatively}
T:  [yeah
      ((T nods several times at S))

T holds clarinet in front of her, gaze resting on S
playing, mouthpiece remains several inches below her
mouth

S:  ^4_3_2_1 {continues third octave, resumes tempo
and fluidity}
S:  =^7_6_5_4_3_2_1 {second octave}
    =^7_6_5_4_3_2_1 {first octave} =1 ____ {last note}
```

Figure 6: The student continues after encouragement

In this short vignette we see the student's use of gaze to seek feedback whilst playing. When this is unacknowledged, it is escalated to a verbal utterance. When gaze is used to seek a response again but encouragement is received, the student continues to play without seeking feedback verbally. We then see how the student uses tempo to express confidence in the descent, holding the turn long enough to complete the scale. Speeding up talk has been shown to be a way to hold on to a turn (Button, 1993).

Delayed interruption

We have previously noted that in a music lesson, where a verbal request by the tutor is usually followed by a musical response, it is reasonable to assume that this musical phrase is analogous to a conversational turn, and we should therefore be able to see the characteristics of turn management (Duffy & Healey, 2012). The tutor will often interrupt the student's playing once a problem has been detected. However rather than stop them as soon as the error has occurred, the end of a musical phrase is preferred by the tutor as a TRP to take the turn, even if they have detected the problem earlier. During the short period of time between detection and interruption, the tutor's non-verbal behaviour reveals their intention to interrupt, such as moving in closer to the music, raising their

arms from their listening pose or picking up a pencil (to write an instruction on the score). If the tutor decides to demonstrate the fragment, they start to pick up their instrument, or bring it closer to a playing position, whilst the student finishes the phrase. Student and tutor will not deliberately play at the same time during this part of the lesson, however we will see later that brief overlap does occur.

We will now examine this in more detail using the notation devised. In this extract, student and tutor have been working on an exam piece together for several weeks (Clarinet Sonata in Eb Major Op.167 Mvmt IV Molto Allegro by Saint-Saens). It includes passages of ascending and descending scales and arpeggios which are challenging to play fluidly. The tutor suggests that they pick up where they left off the previous week. The student starts by playing the phrase shown in Figure 7 however he plays a wrong note from a different scale which changes the tonality of the passage. The tutor indicates that she has noticed the error by adjusting her gaze and listening position but does not interrupt the student yet. He restarts mid-phrase, from just before the error (Figure 7). This may be for the benefit of the tutor, who has to solve a continuation problem, i.e., how to relate the repair to the original utterance (Levelt, 1983) or for his own benefit in ease of correction. The choice of where to start a musical self-repair is another area worthy of further investigation.



Figure 7: Bars 118-120 Clarinet Sonata in Eb Major Mvmt IV Publisher Durand, 1921. Plate D. & F. 10,063, Paris.

The tutor moves closer to the score with her pencil and as the student reaches the long note at the end of the phrase, she talks over it and moves the pencil towards the score (Figure 8). We see that rather than interrupt mid-phrase, she lets the student attempt self-repair, only taking over the turn at the TRP presented by the long note at the end of the phrase.

```
S:  ^ _ _ _ _ _ ^ _ _ _ _ _ * {*error, wrong note}

T turns head up from music to the student, moves in from
her listening position but does not say anything.

S blows into the clarinet and restarts from just before
the error, playing more slowly.

S:  [ ^ _ _ _ _ ^ _ ] [ ^ _ _ _ _ ]
      [T picks up pencil] [T moves into score with pencil]

S:  _____ // _____ (1.37) {final note of phrase}
T:  _____ //OK so make sure that it's A major not E major.
```

Figure 8: The tutor delays interruption

