Hacettepe University Journal of Education (HUJE) 33(Special Issue): 316-333 [2018]
doi: 10.16986/HUJE.2018038809 [Research Article]

Refining Musical Performance through Overlap*

Sam DUFFY**, Patrick G. T. HEALEY***

• ***Received:*** *02.09.2017* • ***Accepted:*** *26.03.2018* • ***Published:*** *25.04.2018*

ABSTRACT: Whilst the focus of attention in an instrumental music lesson is refinement of the student's musical performance, conversation plays an essential role; not just as a way to analyse the student's musical contributions, but to organise them within the lesson flow. Participants may respond to talk through performance and vice versa, or even spend periods of time exchanging purely musical contributions. The short musical fragments exchanged by the participants are managed within lesson dialogue in ways analogous to conversational turn-taking. Problems in the student's performance are refined through both student self-initiated and tutor other-initiated repair, initiated by embodied action and play. A fundamental part of turn-taking is managing the transition to a new speaker. The presence of musical contributions allows for additional types of transition, for example from a turn at talk, to a musical contribution. In conversation, there is generally a preference for a short pause at the transition to a new speaker, and overlap tends to be minimised when it occurs. Through detailed qualitative video analysis of a one-to-one clarinet lesson, we find differences in the preferences regarding overlap when purely musical contributions are being exchanged, and that the duration of overlap during these exchanges of fragments of music are significant.
Keywords: music education, conversation analysis, repair, overlap, interaction

1. INTRODUCTION

Fundamental to a one-to-one instrumental music lesson is the social interaction between student and tutor. Intuition suggests that the principal activity would be playing music, however conversation is just as important, not just as a way to analyse the student's musical contributions, but to organise them within the lesson flow (Duffy & Healey, 2014). Activities are managed conversationally. Lesson discussion is interleaved with performance, demonstration and musical experimentation, resulting in a rich multi-modal interaction. Participants may respond to talk through performance and vice versa, or even spend periods of time exchanging purely musical contributions. These musical contributions could be performances of whole sections of a piece, phrases, or unscripted exchanges of short musical fragments intertwined with the lesson dialogue. Analysis of the shape and timing of these musical fragments shows that they are managed in ways analogous to conversational turn-taking (Duffy & Healey, 2013). Problems in the student's performance are refined through both student self-initiated and tutor other-initiated repair (Duffy, 2015, Chapter 13). A key component of conversational turn-taking is how the participants manage the transition from one speaker to another. There is a preference to minimise overlap. The presence of musical fragments in the dialogue introduces the potential for transitions between turns at talk and musical contributions. When transitions are between two musical contributions, the preference for avoiding overlap does not operate in the same way (Duffy, 2015, Chapter 12). In this article we will first define repair and the types of transition to a new speaker commonly found in conversation, as described in the literature. We will then briefly outline the notion of 'error' in musical performance, and how repair can be applied to the refinement of problems in musical performance. Then, we will work through a detailed example from a one-to-one clarinet lesson, using conversation analysis to illustrate how musical contributions, and specifically overlap, contribute to refinement of the student's performance of a problematic musical phrase. Finally we will review and discuss the findings.

* This data that this article is based on forms part of the lead author's doctoral thesis (Duffy, 2015).

** Dr., Cognitive Science Research Group, School of Electronic Engineering and Computer Science, Queen Mary University of London, UK. e-mail: s.duffy@qmul.ac.uk (ORCID: 0000-0003-4963-0479)
*** Professor, Cognitive Science Research Group, School of Electronic Engineering and Computer Science, Queen Mary University of London, UK. e-mail: p.healey@qmul.ac.uk (ORCID: 0000-0003-3079-3374)

2. REPAIR AND OVERLAP

2.1 Repair

An important part of maintaining conversation is dealing with turn-taking errors or rule violations, such as the misunderstandings or interruptions frequently found in natural dialogue. The turn-taking system provides a locally structured interactional space for detecting and dealing with these problems, known as repair. Note that an 'error' in this context relates to how transition to a new speaker is achieved in relation to commonly understood rules of conversational turn-taking, and is not limited to the content of the utterance itself.

> *The term 'correction' is commonly understood to refer to the replacement of an 'error' or 'mistake' by what is 'correct'. The phenomena we are addressing, however, are neither contingent upon error, nor limited to replacement.* (Schegloff, Jefferson, & Sacks, 1977, p. 363)

It could be that there has been an incorrect assumption of contextual knowledge, or simply that the utterance is ambiguous or vague and further clarification is required. Repair may be initiated by the speaker (known as *self-initiated repair*) or by another party to the conversation (known as *other-initiated repair*). Schegloff et al. (1977) made a comprehensive study of self-repair, noting that it occurred much more frequently than other-initiated repair in natural speech; partly because opportunities for self-initiation come before opportunities for other-initiation, and partly due to a preference by participants for self-repair over other-repair. Even if self-repair is other initiated, i.e. the listener initiates repair in their next turn, the listener is still more likely to encourage the original speaker to correct themselves, than to make a direct correction.

2.1.1 Self-Initiated Repair

Self-initiated repairs may be initiated in one of three main positions (Schegloff et al., 1977, p. 366):

1. within the same turn as their trouble source,
2. in that turn's transition space, or
3. in the third turn to the trouble-source turn, i.e. in the turn subsequent to that which follows the trouble-source turn.

The most common type of self-repair is the first described above: self-initiated-self-repair (SISR), which occurs when the current speaker detects and addresses a problem before the turn in which it occurs is completed. To detect the need for self-repair, the speaker must be able to self-monitor and identify a problem in what they are saying, or see some outward sign of the listener's confusion. 'Correction' is made promptly once a problem has been detected. A neutral holding term such as 'uh' can be used to communicate error detection to the listener, and so hold the turn for the repair to be made. Consider the following example, from a corpus of 959 spontaneous self-repairs made by participants asked to describe visual patterns made up of connected, coloured dots:

"go from left again to, uh..., from pink again to blue" (Levelt, 1983, p. 44).

Levelt explains that such a repair typically consists of three parts; *"go from left again to"* is the original utterance and contains the trouble spot or reparandum which needs to be repaired. The second part is the editing phase, a period of hesitation shown here by *"uh..."*. The third part is the repair proper *"from pink again to blue"*. A repair can start directly at the reparandum, or the speaker can retrace to an earlier point in the turn.

The ‘position 3’ or ‘third turn’ self-repair occurs when the listener’s response indicates to the speaker that they have been misunderstood. The speaker both detects the problem and makes repair, so it is still self-initiated, but only because they have been able to detect misunderstanding from the response to their first (or ‘position 1’) turn (Schegloff, 1992).

2.1.1 Other-Initiated Repair

Other-initiated repair tends to occupy just one main position, the turn subsequent to the trouble-source turn (Schegloff et al., 1977, p. 367), both alerting the speaker that there is a misunderstanding or problem with their turn, and inviting them to address it in their next turn. The most common way that others deal with trouble in speaking, hearing or understanding is to initiate a process of repair with something like “sorry?” or “what?”, but leaving the opportunity to provide the repair solution to the speaker of the trouble source (Sidnell & Stivers, 2013, p. 249). For example:

1. *Lori: But y'know single beds'r awfully thin tuh sleep on.*
2. *Sam: What?*
3. *Lori: Single beds. They're-*
4. *Ellen:→ Y'mean narrow ?*
5. *Lori: They're awfully narrow yeah*

(Schegloff et al., 1977, p. 378)

Repair initiations can be accompanied by gestures, and in some pedagogical contexts they can initiate repair without verbal accompaniments (Sidnell & Stivers, 2013, p. 251). We will see an example later, in section 5.1, where the tutor’s elevation of her clarinet initiates student repair.

2.2 Transition to a New Speaker

Overwhelmingly in conversation just one person speaks at a time (Sacks, Schegloff, & Jefferson, 1974). There is a preference for avoiding gaps and overlaps at the transition between speakers, and whilst instances of overlap occur frequently, they are usually brief. It is possible for a listener to take the floor almost instantaneously at the end of the previous speaker’s turn with no interval between the end of the prior and start of the next piece of talk, known as a latched turn, however this is not a frequently occurring transition type. An intention to take the floor in this way can lead to unintended overlap, for example where a recipient anticipates a possible turn completion and makes a bid for the floor whilst the current speaker continues to talk (Jefferson, 1986). This highlights a challenge for the listener; to be able to accurately predict when the speaker’s turn will end, known as the turn completion point (TCP) in order to take their turn at talk. The listener must process what they are hearing well enough to predict both the content and its structure, so that they can predict when the utterance may come to an end (Levinson, 2012).

It has been proposed that both syntax (Sacks et al., 1974) and prosody (Couper-Kuhlen & Selting, 1996; B. Wells & Macfarlane, 1998) are used to project the end of a speaker’s turn. Speech rhythm is also thought to play a part (Schaller & Müller, 2013). It has been suggested that the lexicosyntactic cues, i.e. the symbolic content of an utterance, are more important than the intonational or prosodic cues, perhaps because prosodic cues occur too late in the turn for forecasting purposes (Ruiter, Mitterer, & Enfield, 2006). However, this finding was based on analysis of data from telephone calls, and the gestural and positional cues available in face-to-face interaction may also play a part. For example, two further explanations for variation in turn transition speed are non-verbal behaviour such as nodding and gaze (Stivers et al., 2009), which would not be available to participants in a phone conversation. Both Stivers et al. and Ruiter et al. agree however on the importance of getting the timing at speaker transition right, since deviation may cause your interlocutor to infer subtleties you hadn’t intended such as impatience or hesitancy.

Rather than the next speaker commencing their turn without any perceivable gap in the conversation, it is much more likely there will be a small pause. Sacks et al. (1974) defined the notion of the transition space, an inter-turn silence which is not treated by the participants as a gap or a pause. This was expanded by Jefferson (1986) as follows:

> *A recipient/next speaker produces his talk in such a way that it occurs with neither haste nor delay. It is not pushed up into, or latched immediately onto the prior utterance, but permits just a bit of space between the end of the prior and the start of the next. I have the impression that this is the most common, the usual, the standard relationship of one utterance to another. (Jefferson, 1986, p. 162)*

Jefferson suggested that recipients are not only monitoring for possible completion of an utterance, but also for a speaker's projected silence thereafter. The definition of turn 'completion' could be further refined to turn 'entitlement' - sensitivity to the speaker's intention to stop or keep going, taking into account the possibility of this transition space. Empirically it has been found that in English, a gap of around 200ms is most likely between turns (Heldner & Edlund, 2010). This is consistent with the findings in a study carried out across a worldwide sample of ten languages (Stivers et al., 2009).

Whilst a short pause is most likely at the transition between one speaker to another, overlap does occur. For example, when more than one person self-selects to take the floor in multiparty conversation, there is a simultaneous start which is resolved by one or other party dropping out (Sacks et al., 1974, pp. 706–708). At natural 'completion points', a recipient may assume that an utterance is complete and bid for the floor whilst the current speaker, perfectly within their rights, continues with their turn (Jefferson, 1986). This can happen when optional additions are placed at the turn completion point, for example, tag questions such as '*isn't it?*', or terms of address and etiquette, for example, '*what's yer name again please sir?'* (Sacks et al., 1974, pp. 707–708). These types of overlap are common and unproblematic, however not all overlap is caused by incorrect judgements regarding when a speaker is approaching the end of their turn. Jefferson (1973) proposes that the occurrence of overlapped tag-positioned address terms is not a trivial misplacement in startings and stoppings, but an intensely organised phenomenon.

Overlap is not always equally distributed between the participants. In a study of dyadic student-tutor meetings, Farr (2003) found that the total instances of overlap and interruption were equally distributed between student and tutor, but that the outcome differed by participant. In all cases an overlap initiated by the tutor resulted in taking control of the floor, whereas not every student initiated overlap resulted in a speaker change. In a study of dyadic student-tutor meetings taking place during the academic 'office hour' in a university, Limberg (2010, p. 181) found that teachers interrupted student presentations to specify certain points or check understanding. Despite this talk being set up as student-managed activity, the tutor's interruptions were seen as implicitly authorised and institutionally sanctioned. The insertions did not lead to an abrupt end of the student's presentation, which simply resumed after clarification had been provided.

In the context of one-to-one music lessons, the presence of musical, as well as verbal contributions allows for four types of transition to a new 'speaker'.

1. Talk following talk.
2. Talk following play.
3. Play following talk.
4. Play following play.

This provides opportunities for different types of overlap. Detailed analysis of a one-to-one clarinet lesson found that these types of overlap were not equally distributed between participants (Duffy, 2015, Chapter 12). For example, the student showed a preference to play over tutor talk whilst the tutor showed a preference to talk over student play. The tutor was often observed to talk over the end of the student's current musical phrase in order to bid for the floor to provide feedback (Duffy & Healey, 2013).

3. REPAIR IN MUSIC PERFORMANCE

A three-part sequence of teacher Initiation, student Response and teacher Evaluation (known as IRE) is a commonly found classroom discourse pattern (Cazden, 2001). The sequence shows the organisation of "questions with known answers" based on the premise that teachers often ask questions whose answers they already know (Macbeth, 2004). However, as is the case for conversation generally, not every problem that arises in pedagogical dialogue is caused by a clearly identified and mutually agreed 'mistake'. In the production of a musical contribution, whilst some problems are technically clear, for example a note played which is different to that indicated in the score, others are more subjective, for example transcript 4.1, where the problem described by the tutor is that the tone feels 'tight' or 'anxious'.

1. T: I think when you (.) just spend a week or two just playing it I
   think you'll find that you feel more comfortable with [bits]
2. S: [hmm]
3. T: of it (.hhh) let's just (0.5s) u::m
4. S: I found with my tuning it went a little bit strange sometimes (0.8s) [I just]
5. T: [uh-hu]
6. T: I-I-I still feel a bit tight
7. S: yea:h
8. T: as though you're being a bit anxious and therefore everything went e:::e
9. S: yea[:h figure 1a.
10. T: [rather than o:[:o
11. S: figure 1b. [my breathing wasn't as (.) good

Transcript 4.1. (Duffy, 2015, p. 108)

Figure 1. T: (a) therefore everything went e:::e (b) rather than o::o (Transcript 4.1 lines 8 and 10)

Other subjective areas could include tempo or volume. Simple tempo problems include playing a passage too fast or too slow, but more subtle decisions concern the rate of change of an *accelerando* (speeding up) or *rallentando/ritardando* (slowing down). The composer will indicate where these should start through markings in the score, and often mark the point by when the desired speed should have been reached, but they do not always mark the precise time that the

player should take to make this tempo change. The rate or shape of change is also often not indicated. This allows for the subjective decision making that contributes to musicianship and expressiveness. Similarly, the precise duration of a pause, or the relative volume change required to make a *diminuendo* (getting quieter) or *crescendo* (getting louder) are areas for the musician to apply their own judgement.

In a detailed analysis of a number of one-to-one instrumental clarinet lessons (Duffy, 2015, Chapter 8), problems in the student's playing of a specific passage were made explicit by the response of the participants. Both student and tutor responded to errors in the student's performance, with different outcomes. The student was observed to monitor the tutor for evidence that an error had been detected as they played, which could then lead the student to interrupt their own performance. Some errors were ignored by one or both of the participants (but not necessarily unnoticed), presumably for the purpose of prioritising and making best use of limited lesson time. Through analysis of the events directly preceding, and following, problems identified by participants during their lessons (Duffy, 2015, Chapter 13), it was possible to find several different ways in which 'musical repair' was initiated. Student self-repair was commonly observed; often self-initiated and made in the same turn as that in which the problem occurred (position 1 SISR). However sometimes the student sought visible tutor acknowledgment of the problem that they had already identified through self-monitoring. By withholding the sought after response, the tutor could initiate a transition space repair (Duffy & Healey, 2013). In other cases, student self-repair was initiated by the tutor's actions as they listened, such as changing posture or altering their grip of tools such as a pencil, or their instrument. The student appeared to interpret the tutor's actions as preparation to bid for the floor to give feedback. Often they were followed by the tutor gesturing over the music, writing on the music, or raising their instrument to demonstrate an alternative phrasing. Whilst these actions are not formally turns at talk, and so do not technically constitute a NTRI, they provide evidence for the student that an error has been detected. They were often followed by student correction of the error that they believed they had made, the tutor choosing not to take to the floor if the correction was satisfactory, allowing the student to continue playing.

Some studies of classroom discourse draw direct analogies between the correction aspect of the IRE sequence and the conversational phenomenon of repair (e.g. McHoul, 1978). However, since not every problem that arises is caused by a clearly identified and mutually agreed 'mistake', it is also proposed that rather than treating the work of repair and classroom correction as alternative expressions of a same organisational domain, they should be understood as different organisations that bear on the production of a same sequence. In other words, that correction is a kind of repair, and it is to repair, and not correction, that we should turn to understand the organisational filaments of both (Macbeth, 2004). It is this approach that is taken in the work that follows.

4. METHOD

Conversation analysis can be used to reveal much about the sensitivities of teaching interaction. For example, there is a large body of work on classroom discourse (e.g. Lerner, 1995; Macbeth, 2004; McHoul, 1978; G. Wells, 1993) and more recently on music tuition (Nishizaka, 2006; Reed & Reed, 2014; Szczepek Reed, Reed, & Haddon, 2013). As part of the data set for a doctoral thesis (Duffy, 2015, Chapter 6), a one-to-one clarinet lesson was observed and filmed. The lesson featured a male student studying ABRSM grade 8 clarinet performance and was recorded during one of his regular weekly lessons at the junior school of a London Conservatoire. The female tutor had taught the student for many years. The student was largely comfortable with the technical challenges of his instrument and capable of exploring musicality and expression. The video data was analysed using the open-source video transcription software ELAN (Brugman, 2004) and qualitative video analysis techniques, as described in Heath, Hindmarsh, & Luff (2010).

Established notation for conversation analysis, as described in the appendix of Sacks et al. (1974), was adapted to analyse the musical contributions to lesson dialogue (Table 1). Specifically, we will use square brackets to show where the musical overlap starts and finishes. A duration inside square brackets e.g. [0.3s] indicates the duration of the overlap.

(0.2s)	Elapsed time (seconds) used to denote pauses or silence
_____(1.4s)	Long single note and duration
_ _ _ (2.3s)	Individual notes in a musical phrase and phrase duration
↑___	Rising passage of notes
↓___	Falling passage of notes
' ' ' (1.2s)	In-breath in preparation to play, and duration
//	Onset of 'talk over play' overlap
{first octave}	Additional information for music notation
[0.6s]	Duration of period of overlap

Table 1: Transcription notation to accommodate music contributions to lesson dialogue (Duffy & Healey, 2013)

5. RESULTS

5.1 Self-Repair Initiated by Embodied Action

This vignette introduces a problem in the student's playing of a phrase, which reoccurs several times during the lesson in which it occurs. This sets up the analysis presented in sections 5.2 and 5.3, where the problems are made explicit and the tutor initiates repair. The student and tutor have been working together for 9 minutes, having discussed exam administration and worked on some scales together. They are now revisiting a piece that they worked on the previous week, picking up from where they left off. The student prepares by taking an in breath and plays the phrase shown in Figure 2.

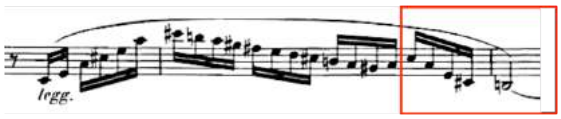

Figure 2. The last five notes of the phrase are a source of trouble in lines 4-6 of transcript 5.1
Bars 118-120 Clarinet Sonata in Eb Major Mvmt IV Publisher Durand, 1921. Plate D. & F. 10,063, Paris

He plays the fourth and fifth notes with a loud squeak (line 1 transcript 5.1) then exhales audibly. The tutor turns her head towards him, raising her free hand to grip her clarinet, perhaps in preparation to bid for the floor through demonstration. However, the student restarts before she has completed raising her hand. The tutor turns her head back to the music, retaining her grip on the clarinet (Figure 3a.) When the student reaches the last five notes in the phrase, he falters, pauses and tries to self-repair, making three attempts (lines 4-6 transcript 5.1).

1. S: ' ' ' ↑_ _ _ {squeak squeak} [((exhales audibly))
[((T moves free hand up to
clarinet, looks at S))
2. S: ↑_ _ _ _ _ _ ↓_ _ _ _ _ _ _ _ _ _| _ _
|[figure a.]
3. S: ↑_ _ {last note incorrect} | (0.4s)
|[figure b.]
4. S: ↓_ _ _ {last two notes incorrect} |(0.4s)
|[figure c.]
5. S: ↓_ _ _ {last two notes incorrect} (0.2s)

6. S: ↓_ _ _ _ ____(0.7s) | (0.2s)
|[figure d.]
[((S lowers
(0.4s) clarinet))]
7. T: ↓_ _[_ _ _____(0.8s)] [___(0.3s)]
8. S: [((S raises clarinet))] [↓_]_ _ _ ____(0.7s)

Transcript 5.1. The tutor changes her grip on her clarinet in response to each of the student's errors - shown by the reference to Figure 3.

Figure 3. The tutor raises her clarinet in discrete stages with each self-repair by the student (see transcript 5.2.)

During the student's attempts, the tutor raises her clarinet to a playing position in discrete stages, raising the instrument at the first significant pause in the student's playing, and then changing her grip on the clarinet each time the student attempts the fragment. Her actions bring the clarinet closer to playing position with each attempt, suggesting that she is preparing to bid for the floor through play. Her actions are escalated in stages, rather than bid for the floor straight

away. Finally, in line 6, the student plays the five notes correctly and finishes the phrase. He sounds the last note for 0.7 seconds then quickly lowers his clarinet from his mouth (Figure 3d). After a brief pause (0.4s), during which time she raises her clarinet to her mouth, the tutor plays. The tutor may not have prepared to bid for the floor earlier as she was waiting to see if the student could self-repair. She therefore needs this brief pause to raise her instrument. She plays the phrase (line 7 of transcript 5.1). As she plays, the student holds his clarinet away from his mouth, raising it again as she plays to take the floor over her last note (line 8 of transcript 5.1).

This example illustrates the tutor encouraging repeated student self-repair through staged movements of her clarinet into playing position (rather than raising her clarinet in one movement) whilst withholding verbal response to a problem. The initial movements are not interpreted as a bid for the floor as they do not end with the tutor in a position ready to play. The lack of verbal interruption by the tutor invites the student to continue to attempt self-repair. It is only once he has played the notes correctly that the tutor makes the final preparation to play, placing the instrument in her mouth. The student does not make a further attempt to play in the pause during which the tutor makes this final movement, recognising this as a bid for the floor, lowering his clarinet away from his mouth.

5.2 Next Turn Repair Initiation Through Play

The next vignette picks up from transcript 5.1. As soon as the student stops playing in line 6, the tutor prepares to play, picking up the short five-note fragment shown in Figure 2, playing it with a lightly 'swung' rhythm (shown continuing from line 7, in transcript 5.2).[1] This provides the student with information on how the tutor would like him to play the phrase next, initiating a further attempt at self-repair through playing. The student starts to play before the tutor finishes sounding her last note (line 8, transcript 5.2). Before the student finishes his last note the tutor plays the phrase again, overlapping with the end of the student's contribution (line 9, transcript 5.2). The student interprets this as feedback that there is still a problem and so makes a second attempt at the five-note phrase (line 10, transcript 5.2). In this way, the tutor uses play to initiate student repair twice, both times through the exchange of musical contributions (Figure 4). The tutor's play effects next–turn-repair-initiation. The tutor keeps her clarinet in her mouth during the student's turns, resting it on her bottom lip, only lifting her top lip to take an in-breath between playing. In contrast, the student briefly lowers his clarinet from his mouth during each of the tutor's turns, further evidence that it is the tutor that controls these exchanges rather than the student.

```
7. T:  =↓ _ _ _ _ _____[ _ ] (1.1s)
8. S:                  [ ↓_ ][0.3s] _ _ _
       _____[ _ ] (0.7s)
9. T:       [ ↓_ ][0.2s] _ _ _ _____[ _ ] (0.7s)
10. S:                               [↓_ _][0.4s] _ _
       _____[ _ ](0.7s)
11. T:      [ ↓_ ][0.4s] _ _↑ _ _ {continues from start of phrase}
```

Transcript 5.2 Student and tutor alternate the five-note fragment.

[1] A rhythm is 'swung' when syncopation is introduced to an otherwise regular, or 'straight' written rhythm. The so-called 'swing ratio' between two successive notes of nominally the same duration is generally taken to be a 2:1 ratio or 'triple feel'(Gabrielsson, 2003, p. 229).

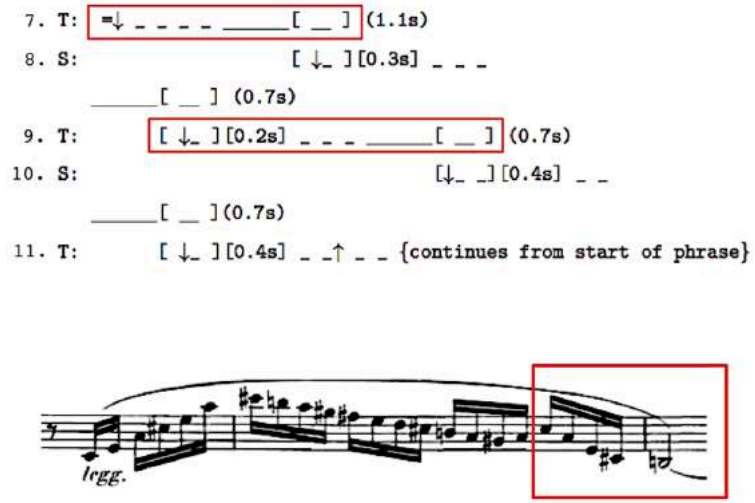

Figure 4. Annotation of transcript 5.2 to show where the tutor uses play to initiate student repair of the five-note fragment.

Before looking at what happens next, we will examine the overlap durations in transcript 5.2. Figure 5 highlights the position of the overlap, at the transition between each turn. It is always at the end of the long note at the end of the five-note phrase, and the overlap duration is reasonably consistent, being between 0.2-0.4s (Figure 6). However, the precise placement of the overlap with respect to the last long note differs.

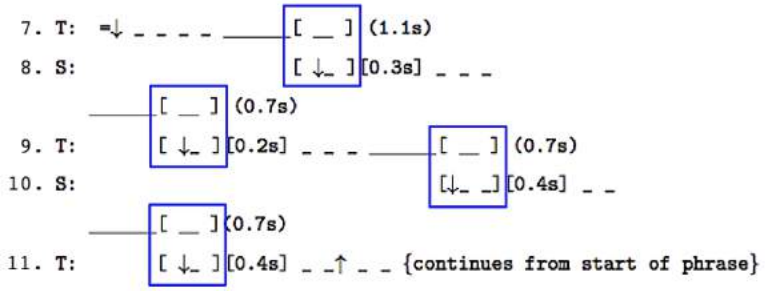

Figure 5. Overlap positions annotated on transcript 5.2.

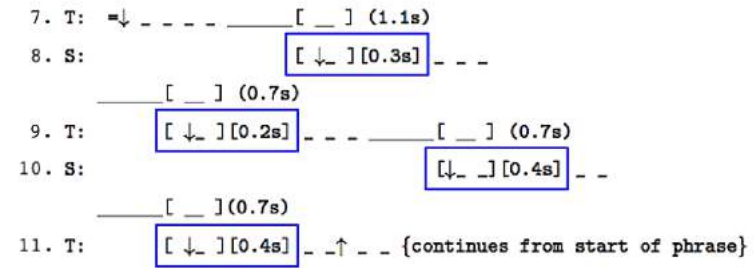

Figure 6. Overlap durations annotated on transcript 5.2.

We see variation in the length of the note played before the other participant interrupts to play the phrase again (Figure 7). When the tutor first played this phrase in line 7, her final note was 1.1s long. We know from having heard both participants play this phrase without interruption that this is the tutor's preferred duration for this note. When the student bids for the floor through

play-over-play overlap in line 8, the tutor continues to play until she has formed the note fully, leading to an overlap of 0.3s. When the student takes his turn in line 8, his last note is only 0.7s long; the tutor interrupted his note after just 0.5s and rather than trying and prolonging the overlap to achieve the preferred note duration, the student drops out after 0.2s of overlap. When the tutor bids for the floor again in line 9, the overlap maintained is slightly longer (0.4s) but she does not play the note for the preferred duration of 1.1s, established previously and reaffirmed in line 7. She drops out to shorten it to 0.7s, the same length as seen in line 8 during the student's turn.

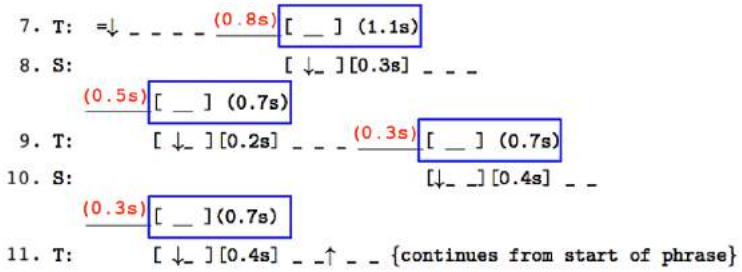

Figure 7. Transcript 5.2 annotated with overlap placement in blue, and final note duration before the onset of overlap in red.

We see several things happening in this excerpt. After the first statement of the shortened five-note phrase, the total duration of the final note is shortened, from the duration of 1.1s seen previously, to 0.7s. The part of this note that is played before the other participant bids for the floor decreases from 0.8s to 0.5s to 0.3s. The duration of the overlap varies between 0.2-0.4s. Both participants are involved in maintaining the duration of overlap, whether they are playing the note, or bidding for the floor over it. The person bidding for the floor continues to play in each case, implying the overlap was intended rather than accidental. So the person playing the note has more responsibility for duration overlap in this case, as they decide when to stop playing to minimise the overlap. After the first statement of the shortened fragment in line 7, both participants are seen to drop out, allowing the final note to be shorter than preferred. This implies that consideration of the phrase in this context has changed; the final note duration is no longer a priority. From this we can infer that the focus of attention is on an earlier part of the fragment.

5.3 Play-Over-Play Overlap Duration

In line 11 of transcript 5.2, the tutor elongated the five-note fragment, adding four more notes at the beginning (Figure 8). Line 11 is shown in full in transcript 5.3, which continues on from transcript 5.2. As seen in line 7 of transcript 5.2, when the tutor presents a new duration of fragment from the phrase for the first time, she plays the last note for the preferred length of 1.1s. The student copies her, also playing the elongated phrase, latching with her last note, rather than making a significant overlap (line 12 of transcript 5.3). As the student approaches the end of the phrase (line 12), the tutor bids for the floor by playing over his last note. The student stops playing after 0.3s of overlap, his last note sounding for 0.8s in total (shorter than the 1.1s played by the tutor in lines 7 and 11). As seen previously, the onset of the tutor's overlap was much sooner than the student's, and he stopped playing rather than sustain the note.

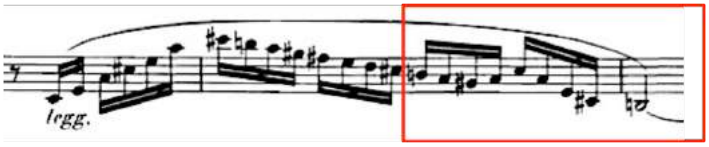

Figure 8. The tutor elongates the fragment.

11. T: ↓_ _ _↑ _ _ ↓_ _ _ _____ (1.1s) =
12. S: =↓_ _ _↑ _ _ ↓_ _ _ _____ [_] (0.8s)
13. T: [↓_] [0.3s] _ _
↑ _ _ ↓_ _ [_ _____(0.7s)]
14. S: [↑_ _ ↓_] ((S grunts)) _ _↓ _ _ _ {*}
(0.3s)
15. S: ↓_ _ _↑ _ _ ↓_ _ _ _____(0.5s){correct ending}

*last note wrong

Transcript 5.3 The tutor elongates the fragment (see Figure 8).

In lines 14 and 15, something happened to interrupt the pattern of alternating contributions - the student started to play over the tutor before she reached the last note of the phrase. The tutor did not drop out after 0.2-0.4s but continued to her final note, although she cut it short to 0.7s. This made the overlap 0.9s in total, and they both stopped playing at the same time. The student grunted and exhaled loudly. He restarted from the beginning of the now nine-note fragment, but made a mistake. The tutor allowed him to continue until he reached the end of the phrase with the correct ending (line 15), but kept her clarinet in her mouth. After a brief 0.5s pause she played the whole 23 note phrase at a faster tempo. The process of alternating the phrase through play continued.

In this vignette, we saw that the student did not overlap with the tutor by playing over her last note when she gave a new instruction for the first time. However the tutor played over the student's last note after his first attempt at the new extended fragment. When the student played over the tutor's last note briefly in line 4 of transcript 5.3, the tutor did not sustain her last note beyond 0.7s, stopping short of her preferred 1.1s. Later, the student started to play before the tutor had reached her last note (line 14 transcript 5.3), but she still played it, albeit cut short of the preferred duration of 1.1s, leading to a period of overlap of nearly a second. The student stopped and grunted, and it was not clear if this was because he had made a mistake or had been affected by the sustained overlap. However, it was noticeable that when the established conventions were breached, regarding the period of overlap acceptable during exchanges of musical contributions, the outcome was not a smooth transition to the next turn.

5.4 'Play-over-play' Overlap Used to Specify a Problem

In the following example, we will again see 'play-over-play' overlap used to refine the size of the musical fragment being played by the student, making it shorter in order to focus on a particular problem. The example starts with the student playing the full phrase that has been worked on in detail earlier in the lesson (Figure 2). As shown in transcript 5.1 and Figure 3, we will see the tutor start to raise her instrument in preparation to play, initiated by a problem with the student's performance.

The example starts 11 minutes into the lesson, and the pair have continued to work on the phrase analysed in transcript 5.1. As the student starts to play, (line 1, transcript 5.4), the tutor listens with her clarinet in her hand (Figure 9a). A squeak occurs at the initiation of a note towards the end of the phrase (shown by *). This specific problem has occurred earlier in the lesson (for example see transcript 5.1). The tutor moves her hand on her instrument as soon as the squeak sounds (Figure 9b) and raises her clarinet to her mouth as the student plays (Figure 9c). As soon as the student finishes the last note of the phrase, he lowers the clarinet from his mouth, indicating that he is not going to continue with the next phrase. He plays the last note for 0.6s (earlier in the

lesson this note had a duration of 1.1s). The student's clarinet leaves his mouth at the same time as the tutor's clarinet reaches hers (Figure 9d). The tutor starts to play, repeating the second half of the phrase where the trouble occurred (line 2, transcript 5.4). By raising her clarinet at the same time as the student plays, the tutor is able to start her demonstration immediately, as soon as the student has finished the phrase.

In line 3 the student takes the floor again, starting from the same point mid-phrase as the tutor did, playing over her last note. He starts to play as soon as the tutor's last note is started. The tutor has shown a preference to play this note for 1.1s previously in the lesson, but here she stops playing after just 0.6s, minimising the overlap caused by the student's play (overlap duration shown by the square brackets in line 3 [0.6s]). The squeak is much less prominent this time but the student cuts his final note even shorter this time, and the tutor takes the floor without overlap in line 4 to play the phrase again. The lack of overlap may be due to the unexpectedly short duration of the student's final note in this instance. This time the tutor plays with an exaggerated swung rhythm and sounds the last note for the full preferred length of 1.1s. She plays 0.7s of this note before the student starts his first note and there is just 0.4s of overlap this time, rather than the 0,6s seen in the previous turn. The student plays the phrase more slowly this time, without the dotted rhythm just demonstrated. He makes the original error from line 1 again (the squeak), but decides to self-repair, replaying the notes preceding the problem note and continuing to the end of the phrase.

Fig 9a. Fig 9b. Fig 9c. Fig 9d.

1. S: ↑_ _ _ _ _ _↓_ _ _ _ _ _ _ _ _ _ _↑ _ _ ↓*_ _ _ _ _____ ((S lowers clarinet))
   {*squeak at beginning of note}
   [((T raises clarinet))]
2. T: _ _↑ _ _ ↓_ _ _ _ [_____](0.6s)
3. S: [_ _↑ _ _][0.6s] ↓_ _ _ _ _ ((last note short))
4. T: _ _↑ _ _ ↓_ _ _ _ ____[_](1.1s) ((T uses exaggerated dotted rhythm))
5. S: [_][0.4s] _↑ _ _ ↓*__ (0.4s) _ _↑ _ _ ↓_ _ _ _ _____
   ((more slowly without the dotted rhythm demonstrated))
6. T: {_ _ ↓_ } {_ _ ↓_ } [{_ _ ↓_ } {_ _ ↓_ } {_ _ ↓_ }]
7. S: [{_ _ ↓_ } {_ _ ↓_ } {_ _ ↓_ }] [{_ _ ↓_ }]
8. T: [Now(0.5s)]=
9. T: =when you-when you go from G sharp to A make sure you're
   sti::ll (0.5s) on the side of your finger I think you're doing that
   and then you've gotta slide back to the G sharp

Transcript 5.4. The student plays a phrase and the tutor responds through play (figure labels and underlines indicate the position of figures 4a-d in line 1 of the transcript).

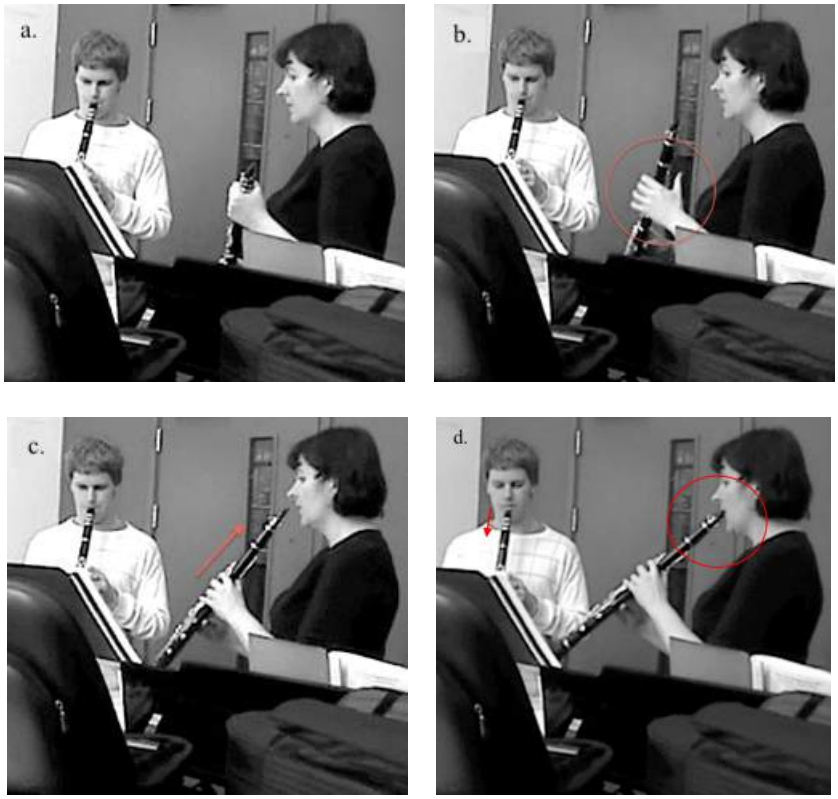

Figure 9. The tutor raises her clarinet as the student finishes the phrase in line 1 of transcript 5.1.

The tutor does not start to play over his final note this time. When she does play (line 6 of transcript 5.4) she introduces a new, even shorter fragment, cutting it down to just three notes, the third being the problem note. She plays the notes repeatedly with a swung rhythm, and a strong accent on the problem note. It appears that the focus of her attention has been drawn back to the squeak, rather than the rhythmic emphasis of the playing. After the first two repetitions the student joins in. The tutor does not drop out but carries on playing with the student for three further repetitions, overlapping for 2.5s. She then drops out to talk (line 8). Her first word "Now" overlaps with the student's fourth repetition. She pauses for 0.5s whilst the student finishes the fragment he has started and then continues her talk. The student now stops playing. The tutor's explanation in line 9 reveals that she has diagnosed the problem causing the squeak (related to not switching between fingerings quickly enough).

6. DISCUSSION

The specific phrase that is the subject of this analysis (Figure 2) was determined by where the tutor asked the student to play from when they began work on the piece, and where he then first experienced difficulty. Student and tutor exchanged musical contributions based on this phrase, as described in transcripts 5.1-5.4 above, for nearly a minute. In the first section, transcript 5.1, we saw the tutor initiate student self-repair through staged movements of her clarinet into playing position, withholding verbal response. In transcript 5.2 the participants moved into a section exchanging a short fragment of the phrase, the tutor initiating a change in the student's approach through play. The final note of the fragment was shortened and the turn transitions were characterised by short overlaps. Having established a preferred duration for the final note of the phrase in earlier parts of the lesson, we saw precise placement of the onset of the overlap at turn transitions. In the next section, transcript 5.3, the tutor elongated the fragment, again restating the preferred duration for the final note, before it was shortened as the fragment was exchanged. In

transcript 5.4, the tutor led the student into playing shorter and shorter fragments through demonstration, focussing on the area containing the recurring problem. After a period of short overlapping phrases, the student converged with the tutor and they played a fragment together, until the tutor withdrew, leaving the student playing alone. In this way, the tutor refined both the length of the fragment to be worked on, and the student's performance, purely through play. The tutor also demonstrated to the student how a difficult phrase could be broken down into smaller fragments, and then put back together when they have been worked on individually. Additionally, by initially not responding to the squeak in line 1 of transcript 5.1, the tutor encouraged self-repair and prioritised the part of the phrase to be worked on. A specific diagnosis and proposed verbal solution to the squeak was not provided until line 9 of transcript 5.4. It is interesting to note that the duration of overlap during the exchange of musical contributions was reasonably consistent at 0.2-0.4s. Where there was not overlap, it was due to the student finishing his last note unexpectedly early (line 3, transcript 5.4). The participants' shortening of the duration of their final note, once the exchange had started, is analogous to the preference for overlap to be minimised during conversation, achieved by a party dropping out.

In speech, the listener is required to project the speaker's utterance in order to plan their bid for the floor (Levinson, 2012). This is a complex task and it has been suggested that syntax, prosody and speech rhythm all play a part in determining when there might be an opportunity to bid for the floor. In conversation "what parties say is not specified in advance" (Sacks et al., 1974, p. 710), but when notated musical phrases are being exchanged, there is potential for each participant to have some knowledge in advance of the shape, construction and duration of the unfolding turn. It is this that makes some of the phenomenon analysed here possible. However it was noted that whilst both participants could assess how closely the student's performance matched the notes written on the score, and both participants had the same opportunity to perceive unmusical sounds or wrong notes, such as the squeak in line 1 or the wrong note played in line 3 and 4 of transcript 5.1, it was the tutor who decided which of these two problems to work on, initially ignoring a squeak earlier in the phrase (line 1, transcript 5.1) and focussing attention on the last five notes (line 7, transcript 5.1). The squeak is not fully addressed until later in transcript 5.4. Whilst there is a preference for the student to self-repair problems in their musical contributions, the tutor prioritises which problems are to be worked on in more detail. This is analogous with turn-taking asymmetries reported in classroom dialogue (McHoul, 1978).

The pitch of each note in the phrase is also predetermined by the musical notation, making analogies with the use of prosody to project the end of a turn difficult in this context. However there is still potential for differences in how each note is reproduced, each time the phrase is played. The microstructure of musical performance goes beyond the musical elements indicated in the score (Gabrielsson & Juslin, 1996), this is how a player is able to differentiate their performance of a set piece from another player, or convey expression. For example, when playing a specific note pitch, it is possible on some instruments, such as wind, brass and strings, to 'bend' and shape notes through micro adjustments in tuning known as intonation. So whilst the fundamental frequency of the note in the musical score is established, something analogous to prosodic variation in speech may still be present.

It was noted that the duration of the final note was often performed for longer than mathematically correct in relation to the other notes, as indicated by the musical notation. Performers on a wide variety of musical instruments have been observed to use phrase-final lengthening at boundaries by increasingly lengthening successive tones as they approach a structural boundary, even when those tones are notated as equivalent duration in the musical composition (Palmer & Hutchins, 2006). Once the duration of the final note in the phrase had been established and agreed by the participants to be around 1.1s, disruption (shortening) of this was

significant. The musical notation also predetermined rhythm, yet the tutor was observed to introduce a swung aspect to part of the phrase, which the student then replicated.

Refinement of a student's sound production is important in another pedagogical context, foreign language tuition. Repair is also an important feature of the organisation of conversation in language lessons. In additional to the conversational turn-taking which occurs in order to maintain lesson flow, the language teacher will specifically seek to initiate repair for problems identified with the student's production of utterances in the language being learnt. As well as the problems with syntax and vocabulary this will include prosodical refinement of pronunciation, cadences and accent. Apart from teachers' correction of learners' utterances, other correction-like activity also occurs, for example learners modify their own utterances in order to correct what seems erroneous to them. This might be self-initiated or initiated by an action of the tutor (Kasper, 1985). In the examples analysed here, we saw that the musical fragments that formed part of the lesson were also refined by the student modifying his own performance, and that his self-repair could be pupil or tutor initiated. Whilst analogies can be drawn between learning to perform music and learning to speak a language, there are also differences. In the language classroom, it can be difficult to distinguish between the official activities of understanding and correcting that are central to learning, and the problems of conversational repair that arise in the course of learning. Where language is simultaneously the tool of instruction, and the object of pedagogical attention, the analytic process of "*discriminating the main trajectory of the interaction from temporary suspension of it for repair*" can be especially difficult as "*not every correction is repair, not every problem in understanding implicates the operations of repair for its solution*" (Schegloff, Koshik, Jacob, & Olsher, 2002, pp. 7–8). In a music lesson, the sound being worked on is distinct from the language used to organise and critique the musical contributions. As a result, where the product of the lesson is musical performance there are opportunities for the new forms of pedagogical repair initiation detailed here.

Some of the recognised ways to manage and hold turns are less available to the musician. One way to hold a conversational turn is through speed. A speaker approaching the possible completion of a turn-in-progress may move to interdict another speaker's bid for the floor, by accelerating the pace of the talk and 'rushing through' the transition space into the start of a next turn-constructional unit (Schegloff, 2000). This is more difficult for a student making a musical contribution, since part of the skill they are displaying is their ability to control tempo, and a long single note has no 'speed' as such, however it may still be possible to achieve holding onto a turn in this way through play, due to the manipulation of rhythm and tempo that is part of musicianship and expressive performance. Another characteristic of the end of a verbal turn is a drop in pitch, or the intake of breath (Button, 1987), whilst a 'rush through' is usually characterised by maintaining the pitch and withholding the in-breath, phrasing the talk to bridge what would otherwise be the juncture at the end of a unit (Schegloff, 1982). Sound production with a clarinet relies on breath control consistent with the musical phrasing, so this mechanism is less available. However, as discussed previously, whilst pitch is pre-determined by the musical score there is potential for subtler pitch adjustments or intonation. In less scripted musical performance, such as jazz improvisation, musicians use this microstructure in combination with non-verbal communication to determine when a musical bid for the floor can be made, taking the 'soloing' role from another player (Healey, Leach, & Bryan-Kinns, 2005). This microstructure could also aid in projection of the end of performance of a notated musical phrase, and whilst this was not examined here, this would be an interesting area of further investigation.

ACKNOWLEDGEMENTS

Many thanks to the reviewers for their helpful suggestions to improve this article. This work is supported by the Media and Arts Technology programme, EPSRC Doctoral Training Centre EP/G03723X/1.

7. REFERENCES

Brugman, H. (2004). Annotating multimedia/multi-modal resources with ELAN. In *In Proceedings Fourth International Conference on Language Resources and Evaluation (LREC '04)*. (pp. 2065–2068). Portugal.

Button, G. (1987). Answers as interactional products: Two sequential practices used in job interviews. *Social Psychology Quarterly*, *50*(2), 160–171.

Cazden, C. B. (2001). *Classroom discourse: The language of teaching and learning* (2nd Edition). Portsmouth: Heinemann.

Couper-Kuhlen, E., & Selting, M. (1996). Towards an interactional perspective on prosody and a prosodic perspective on interaction. In Elizabeth Couper-Kuhlen & M. Selting (Eds.), *Prosody in conversation: Interactional studies* (pp. 11–56). Cambridge: Cambridge University Press.

Duffy, S. (2015). *Shaping musical performance through conversation*. Queen Mary University of London.

Duffy, S., & Healey, P. G. T. (2013). Using music as a turn in conversation in a lesson. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society* (pp. 2231–2236). Berlin: Cognitive Science Society.

Duffy, S., & Healey, P. G. T. (2014). The conversational organisation of musical contributions. *Psychology of Music*, *42*(6), 888–893. http://doi.org/10.1177/0305735614545501

Farr, F. (2003). Engaged listenership in spoken academic discourse: the case of student–tutor meetings. *Journal of English for Academic Purposes*, *2*(1), 67–85. http://doi.org/10.1016/S1475-1585(02)00035-8

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music*, *31*(3), 221–272. http://doi.org/10.1177/03057356030313002

Gabrielsson, A., & Juslin, P. N. (1996). Emotional expression in music performance: Between the performer's intention and the listener's experience. *Psychology of Music*, *24*(1), 68–91. http://doi.org/10.1177/0305735696241007

Healey, P. G. T., Leach, J., & Bryan-Kinns, N. (2005). Inter-play: Understanding group music improvisation as a form of everyday interaction. In *Proceedings of Less is More — Simple Computing in an Age of Complexity*. Cambridge: Microsoft Research.

Heath, C., Hindmarsh, J., & Luff, P. (2010). *Video in qualitative research: Analysing social interaction in everyday life*. London: SAGE Publications Ltd.

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, *38*(4), 555–568. http://doi.org/10.1016/j.wocn.2010.08.002

Jefferson, G. (1973). A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences. *Semiotica*, *9*(1), 47–96. http://doi.org/10.1515/semi.1973.9.1.47

Jefferson, G. (1986). Notes on "latency" in overlap onset. *Human Studies*, *9*, 153–183.

Kasper, G. (1985). Repair in foreign language teaching. *Studies in Second Language Acquisition*, *7*(2), 200–215.

Lerner, G. H. (1995). Turn design and the organization of participation in instructional activities. *Discourse Processes*, *19*(1), 111–131. http://doi.org/10.1080/01638539109544907

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41–104.

Levinson, S. C. (2012). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The handbook of conversation analysis* (pp. 103–130). John Wiley & Sons.

Limberg, H. (2010). *The Interactional organization of academic talk: Office hour consultations*. John Benjamins Publishing.

Macbeth, D. (2004). The relevance of repair for classroom correction. *Language in Society*, *33*(5), 703–736. http://doi.org/10.1017/S0047404504045038

McHoul, A. (1978). The organization of turns at formal talk in the classroom. *Language in Society*, *7*(2), 183–213. http://doi.org/10.1017/S0047404500005522

Nishizaka, A. (2006). What to learn: The embodied structure of the environment. *Research on Language & Social Interaction*, *39*(2), 119–154. http://doi.org/10.1207/s15327973rlsi3902_1

Palmer, C., & Hutchins, S. (2006). What is musical prosody? *Psychology of Learning and Motivation - Advances in Research and Theory*, *46*, 245–278. http://doi.org/10.1016/S0079-7421(06)46007-2

Reed, D., & Reed, B. S. (2014). The emergence of learnables in music masterclasses. *Social Semiotics*, *24*(4), 446–467. http://doi.org/10.1080/10350330.2014.929391

Ruiter, J. De, Mitterer, H., & Enfield, N. (2006). Projecting the end of a speakers turn: A cognitive cornerstone of conversation. *Language*, 515–535.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735.

Schaller, F., & Müller, H. (2013). The contribution of speech-rhythm to end-of-utterance detection. In *Proceedings of the 17th Workshop on the Semantics and Pragmatics of Dialogue*.

Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of "uh huh"and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse: Text and talk* (pp. 71–93). Washington: Georgetown University Press. http://doi.org/10.2307/324165

Schegloff, E. A. (1992). Repair after next turn : The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, *97*(5), 1295–1345.

Schegloff, E. A. (2000). Overlapping talk and the organization of turn-taking for conversation. *Language in Society*, *29*(1), 1–63.

Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, *53*(2), 361–382.

Schegloff, E. A., Koshik, I., Jacob, S., & Olsher, D. (2002). Conversation analysis and applied linguistics. *Annual Review of Applied Linguistics*, *22*, 3–31.

Sidnell, J., & Stivers, T. (Eds.). (2013). *The handbook of conversation analysis* (Paperback). West Sussex: Wiley-Blackwell.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10587–92. http://doi.org/10.1073/pnas.0903616106

Szczepek Reed, B., Reed, D., & Haddon, E. (2013). NOW or NOT NOW: Coordinating restarts in the pursuit of learnables in vocal master classes. *Research on Language & Social Interaction*, *46*(1), 22–46. http://doi.org/10.1080/08351813.2013.753714

Wells, B., & Macfarlane, S. (1998). Prosody as an interactional resource: Turn-projection and overlap. *Language and Speech*, *41*(3–4), 265–294. http://doi.org/10.1177/002383099804100403

Wells, G. (1993). Reevaluating the IRF sequence: A proposal for the articulation of theories of activity and discourse for the analysis of teaching and learning in the classroom. *Linguistics and Education*, *5*(1), 1–37. http://doi.org/10.1016/S0898-5898(05)80001-4