

## **SUPPORT AND EVIDENCE FOR CONSIDERING LOCAL CONTINGENCIES IN STUDYING AND TRANSCRIBING SILENCE IN CONVERSATION**

Israel Berger

### **Abstract**

Using a conversation analytic methodology, this report looks at conversations in English in which lengthy silences are regularly present. These silences are treated as unproblematic in this corpus. They apparently deviate from the proposals that gaps are minimized (Sacks, Schegloff, & Jefferson 1974) and that there is a standard maximum silence of one second (Jefferson, 1988). This is discussed in light of context and culture. Then the robustness of some features of the organisation of sequences (Schegloff 2007) and turn-taking (Sacks, Schegloff, & Jefferson 1974) are considered. Finally, solutions are compared for rendering lengthy silences in such a way that their meaning is preserved in conversation analytic transcripts or others that include timed silences.

**Keywords:** Silence; Pause; Gap; Lapse; Conversation analysis; Transcription; Preference; Sequence organisation; Turn taking.

### **1. Introduction**

Sacks Schegloff & Jefferson (1974), in their foundational work on turn taking, show as one of fourteen ‘grossly apparent facts’ for which their model accounts, that silences and overlap are minimised in transition from one speaker to another, i.e. that no gap, no overlap between turns is a consequence of the systematics of turn taking. The time between turns is not presented as a steadfast rule of any specific time, but rather as a distribution with one beat of silence as the norm. This introduces the problematic of preference<sup>1</sup> into turn taking organisation.

---

<sup>1</sup> Preference in conversation analytic literature refers to a pragmatically preferred action, rather than an psychologically preferred action. For example, the preferred response to an invitation is acceptance, not because of any desires on the part of the participants but as a consequence of the structure of conversation. Preferred responses tend to come without delay (silence, perturbations) and without much elaboration. On the other hand, dispreferred responses tend to be delayed and elaborate. See Schegloff (2007) for a detailed treatment of preference in sequence organisation.

After a speaker produces a first pair part, a delay in responding is often taken to indicate that a dispreferred second pair part might be forthcoming. When such a delay is a silence, it can be termed a *gap*. (A response could also be delayed through the production of vocalisations such as *uh*, *um*, and so on.) Such gaps (or delays) may be an issue for the speaker of the first pair part. Since talk at this point is relevant and noticeably absent, speakers can do work, such as adding increments or rephrasing the first pair part, to elicit a preferred response. (cf. Pomerantz 1984). Gaps may be distinguished from *pauses* within a speaker's talk and from *lapses*, where talk (or vocal conduct such as laughter) does not occur and is not conditionally relevant. An example of the latter would be after a sequence has been completed. (cf. Pomerantz 1984)

Jefferson (1988), in an analysis presented as 'preliminary notes', suggests that speakers use a metric that results in speakers rarely allowing silences of more than approximately one second. She refers to this as a 'standard maximum' and shows here that longer silences tend to occur in the context of non-talk activities such as consulting a train timetable or writing directions. However, there have long been claims of cultural variation in silence in conversation (e.g., Goffman 1963: 103).

Stivers et al. (2009) tested the competing claims of universality and cultural variation by measuring inter-turn silence in 10 languages spoken in informal conversation. Although any similarities among 10 languages can hardly be called universal, possible differences can become known even with a relatively small data set. They found a difference of approximately 0.5 sec between the languages with the fastest and slowest mean responses (Japanese 7.29 ms and Danish 468.88 ms) and that the distribution patterns were similar across the 10 languages. Stivers et al. suggest that differences in response timing may be due to cultural differences in the *calibration* of what is considered delayed.

Whilst it is a worthy endeavour to test claims of variation or universality quantitatively, their study possesses potential problems with generalising beyond the specific languages and sequence types they examined. Their study was limited to polar questions, questions that are designed to receive 'yes' or 'no' responses. Because such questions have a relatively small pool of possible responses, this could result in more uniformity (and therefore lead to a false perception of universality). They base the generalisation from questions to all turn types across all languages on a corpus of Dutch data in which response times were not significantly different between questions and non-questions. It is worth noting that Dutch also did not have significant differences on response timing for confirmation/disconfirmation or gaze/no gaze, whereas most other languages did. It is also unclear in what contexts the recordings were made. 'Informal' could mean anything from a couple talking over a meal to a lively gossip session among friends, and context could be a confounding variable making languages appear less similar than they may actually be.

Gardner, Fitzgerald & Mushin (2009) in studying interactions among Garrwa-speaking Indigenous Australian and English-speaking Anglo-Australian interactants, show that there was both a tendency toward lengthy silences (compared to canonical data) as well as an orientation to smooth speaker transitions as explicated in Sacks, Schegloff & Jefferson (1974). Gardner et al. (2009) argue that tolerance for pauses and gaps is dependent on local interactional contingencies, namely a relaxed atmosphere and lack of orientation to 'getting things done'. They maintain that whilst this attitude may be prevalent within a culture, a tendency toward long silences is not necessarily a property of the culture itself. This argument is supported by their analysis of Australian

political debates in which there are quick responses to prior talk and indeed quite a bit of overlap. I will address these issues as they occur in non-Australian English data.

Sacks, Schegloff, and Jefferson used data that was primarily derived from audio recordings for both co-present and telephone conversations in California and the Northeastern United States in densely populated cities with white participants from a variety of socioeconomic backgrounds. The data presented here is from a variety of everyday and institutional interactions in large cities in England and the Mountain and Southeastern regions of the United States, with participants from white, Black<sup>2</sup>, and Jewish backgrounds. We do not propose that particular demographics account for differences per se, but that a diverse data set is necessary for understanding differences, or indeed similarities, among interactions. Additionally, it should be noted that like the Gardner et al. corpora, the data presented in this report are video recordings of co-present interactions. Video recordings potentially allow analysts to attribute more accurate meaning to silences through knowledge of non-verbal behaviour, including gesture and engagement with the environment. These factors may affect whether a participant is taken to be not responding or in the process of responding, but there is no way to know whether they are occurring at all in audio-only data.

I shall provide additional evidence that tolerance for silence is shaped by local interactional contingencies, and indeed may shape or renew those contingencies. I shall show how interactants orient strongly to sequence organisation, speaker selection, and turn organisation despite this apparent deviation of long silences from canonical turn taking procedures. Finally, I will consider some of the issues in transcribing data that contains longer than canonical silences.

## **2. Deviations and support**

### ***2.1. Deviation from canonical turn-taking features***

In the data we present, there are often long silences of varying lengths. Whilst this is not precluded by Sacks, Schegloff, & Jefferson (1974), they do treat the possibility of long silences as a tool for and feature of dispreferred responses and normative turn taking as no gap, no overlap. We find that this is not always the case for normative turn taking, and in fact the presence of silence within and between turns is often attributed to cultural differences in speaking and interaction (e.g., Tannen 1984).

The current research provides support for Gardner et al.'s argument 'against equating an instance of turn allocation that differs from that identified in the SSJ as a deviation tied to a specific culture'. In other words, whilst there may be overall differences in speech rate or conversational rhythm across cultures and languages, this is not an absolute attribute across members of a culture or individual interactions within a culture or set of participants. I will demonstrate with a case of two different styles during the same interaction. The following two extracts are taken from the same

---

<sup>2</sup> 'Black' is used here in recognition of participants' ethnic identities rather than external categorisation denoted by 'black'. It is used with the understanding that others may identify as 'African American' or use other terms. (For further discussion, see McWhorter 2004; Gourley 1973; and This Wicked Day 2010).

psychotherapy session<sup>3</sup>, that is, the same therapist (Paula) and the same client (Leif) on the same day with the same speaker having epistemic authority and the floor, overall. Throughout this report, I will refer to extracts from Paula & Leif's sessions, primarily for uniformity for comparison between extracts but also because this subset of the data contains the lengthiest silences, and therefore the clearest examples, in the corpus.

These extracts are from video-recorded psychotherapy sessions of approximately one hour duration in a large but low population density city in the Mid-western United States. Sessions between Paula and Leif make up approximately six hours of data. They had been seeing each other for three years when data collection began, and the extracts presented are from their first recorded session. Paula was loaned a consumer-grade camcorder by the author for the purpose of recording sessions, and both participants were able to direct when recordings were made and whether the recording of a session should be ended early or deleted. Paula is a humanistic psychotherapist whose training is in social work, and Leif is seeing her for issues related to both recovery from alcoholism and mobility problems. Both are middle-aged Americans, with Leif being somewhat older. They are from different parts of the Mid-western United States and have lived in their current area for several years. This data has been transcribed by the author using Jeffersonian conventions, and silences were timed initially with a stop watch before being checked against the digital wave files extracted from the video. Times for silences represent exact measurements rounded to the nearest 1/10 second. See Section 3 for a discussion of silence transcription methods and relative vs. exact measurement options.

In Extract 1, Leif is talking about himself and his day to day life, and it contains lengthy silences. He designs his turns so that silences are generally before a turn constructional unit (TCU) is possibly complete. Paula does not produce utterances to break these silences, treating them as pauses after which the TCU will be resumed. She does produce utterances after possibly complete TCUs and long silences separate their turns.

(1)

- 1 Leif: >Yeah it'd been kinda an interesting week=I haven't  
 2 gotten this much accomplished< (1.6) as I'd like to  
 3 this week  
 4 (2.8)  
 5 Leif: A:::n (1.1) I decided that maybe that's okay  
 6 (1.6)  
 7 Leif: I'm not gonna (0.7) beat myself up about it  
 8 Leif: I mean nobody's (1.4) nobody ELSE'll beat me up  
 9 about it so: (1.9) why should I be ↑first in li:ne?  
 10 Paula: How long was it that you fired your boss?

---

<sup>3</sup> There have been a number of articles on silence during psychotherapy sessions as part of a therapeutic approach (e.g., Prince 1997). Most of these have focused on the meaning of silences and are based on clinical experience rather than empirical studies. None have looked at silence in psychotherapy as a linguistic device. It does not appear to be the case that the silence discussed here is part of a therapeutic approach but rather is endogenous to the interaction itself. This is evinced by the presence of silence that can be attributed to either the therapist or the client regardless of prior speaker or who has rights or obligations to speak next. Additionally, there are instances of 'doing being silent' or 'doing not responding yet', which the majority of silences between this dyad are not.

11 (1.0)  
 12 Paula: °°(I mean)°°  
 13 (1.7)  
 14 Paula: How many months  
 15 (2.3)  
 16 Leif: Well January third so I guess that'd be:: uhh (1.6)  
 17 you know (0.6) well goin on eight months  
 18 (4.8)  
 19 Paula: 's still sinking in  
 20 (0.6)

Compare this style to that of the next extract. In Extract 2, Leif is telling Paula how to solve a computer problem at the beginning of the session. His turns do not contain as many pauses, and silences are much shorter. There are even some instances of 'precision timing' in speaker transition.

(2)  
 1 Leif: So you bought it about a year ago  
 2 Paula: Hmhm ((nodding))  
 3 Leif: You probably have Leopard  
 4 (0.9)  
 5 Paula: ((nodding)) Y::es (0.7) I do  
 6 Leif: =(It's an/Nice) operating system  
 7 Leif: hh and and (0.6) Leopard (1.5) has the ability to  
 8 recognise (1.2) a Windows file  
 9 (0.8)  
 10 Leif: When it sees one  
 11 (0.7)  
 12 Leif: It goes  
 13 Paula: (okay)  
 14 (0.7)  
 15 Leif: A↑::hhh (1.6) that's one a THOse (.) files  
 16 Leif: hh and so (0.8) a (0.1) a Windows (0.7) M S Word  
 17 file (.) will open up (1.2) err-r show on the screen  
 18 rather=  
 19 Paula: =Mhm  
 20 Leif: =(0.6) as (1.1) y'no a do::t do::c (1.7) file  
 21 (0.5)

Overlap is virtually absent in the corpus. However, there does not seem to be an inverse relationship between lengths of silence and propensity for overlap. What little overlap we encounter tends to occur in contentious situations, such as in Extract 3.

(3)  
 1 Leif: No I don't miss [that  
 2 Paula: [.hhuh  
 3 Paula: You don't miss: (1.1) falling down becau::se of  
 4 drinking

- 5 Leif: W' I didn't usually do that  
 6 Paula: Pardon  
 7 Leif: I didn't usually fall down  
 8 (0.8)  
 9 Paula: Even with the PAWS  
 10 (.)  
 11 Paula: Even when um (0.8) your doctor was talking  
 12 about the post withdrawa[  
 13 Leif: [That that really .hh ['s  
 14 Paula: [That kind of was a time limi[ted thing  
 15 Leif: [That was really ataxia  
 16 (2.2)  
 17 Leif: An so it was kind of a \*clugeon of\* of ((\*brings  
 18 hands together interlocking fingers slightly  
 19 during these words))  
 20 Paula: ↓Yeah  
 21 (1.2)

In this extract, Paula has enquired whether Leif misses drinking, which he has denied. She has challenged this with the argument that drinking is a lot like smoking (she is a smoker), and that some times are perfect moments for a cigarette. She has gone on to list a few things that she proposes he doesn't miss. This extract immediately follows that list. Paula now (at line 3) asserts that Leif doesn't miss falling down *because* of drinking. This is something that Leif has apparently not experienced, and he is quick to correct the apparent assumption. Paula tries to show her knowledge of previous events by introducing withdrawal as a time when Leif fell down (i.e. not because he was drunk, as previously implied). In lines 12-15, there is significant overlap as they compete to establish the correct interpretation of Leif's situation. In lines 11-12, Paula has rephrased her question from line 9. This ends with terminal overlap from Leif's turn at line 13. Leif's turn is then in terminal overlap with Paula's next turn at line 14, which suggests that his falling down was time limited. This turn is then recognitionally overlapped in the middle of 'limited' with Leif's next turn at line 15, 'That was really ataxia'. There is silence for 2.2 seconds, and Leif produces a further explanation of his prior situation during the time to which Paula has referred, that it was a 'clugeon of ((gesture))'.

A gloss of this turn is that the situation was a bundle of different problems put together. At this point, Paula concedes to Leif's epistemic authority about his own medical problem. Although it may appear that Paula is attempting to write Leif's history for him, she has gotten into a tight interactional place with her earlier assertion (as well as the assumption that Leif must miss drinking sometimes) and must choose between explicitly acknowledging that she was wrong or attempting to find some basis for her assertion after the fact. Nevertheless, the participants must establish what will be accepted as fact in the local interactional environment (regardless of what they may actually believe).

In the previous extracts, we did not encounter any clearly dispreferred responses. In Extract 4, Paula enquires as to how Leif is doing (line 2). For 3.5 seconds he maintains his body and head position that he had prior to Paula's question and then looks up and says 'I'm okay' and smiles (line 4). This initial response is delayed and is analysable as

disingenuous (thus being dispreferred in structure albeit with a word choice that would otherwise be a preferred response). Note that the delay is not necessarily in timing alone. Leif holds his position during the gap at line 2. He is *doing* not responding. Paula treats 'I'm okay' as disingenuous and challenges Leif's assertion by asking him to account for his reasoning, which after two attempts results in Leif confirming that he is indeed not okay. He has significantly delayed his dispreferred response in the form of beginning with linguistic elements of a preferred response but with the prosody expected in a dispreferred response.

(4)

- 1 Leif: ((holding hands together and gazing down))  
 2 Paula: So how are ya  
 3 >> (3.5) ((Leif holding position))  
 4 Leif: >> ((looks up)) °I'm okay° ((smiles))  
 5 (1.1)  
 6 Paula: How do you know  
 7 (2.4)  
 8 Leif: I said so:  
 9 (2.3)  
 10 Paula: On your own authority you are okay  
 11 Leif: No  
 12 (1.9)  
 13 Leif: ((sniffs))

In this extract, there are some very long gaps. However, there are equally long gaps in other extracts in which there are no signs of dispreference, for example extracts 5 and 6.

(5)

- 1 Paula: >What did you do<  
 2 (1.4)  
 3 Paula: <I know it was magic  
 4 >> (3.9)  
 5 Leif: Uh::m (2.8) I (1.1) saved (0.6) th' files (3.1) and  
 6 renamed them  
 7 >> (2.8)  
 8 Leif: For instance (0.6) uh (2.1) uh- (4.2) th- (0.3)  
 9 >first one was called uhh< (0.8) oh gee I drew a  
 10 blank uhh reservation er something like that

(6)

- 1 Leif: Well January third so I guess that'd be:: uhh (1.6)  
 2 you know (0.6) well goin on eight months  
 3 >> (4.8)  
 4 Paula: 's still sinking in  
 5 (0.6)  
 6 Paula: There's no one to account to  
 7 (0.6)  
 8 Paula: <↑But your wife

- 9 (1.5)  
 10 Paula: An the dawg  
 11 >> (2.2)  
 12 Leif: ↓Yeah

Communication (formerly Speech) Accommodation Theory (Street & Giles 1982; Giles & Wiemann 1987) as well as Bateson's schismogenesis (1935) state that people change their communication styles to suit the context, and specifically to be in line with their interlocutors' styles. In interactions where hostility and power are exhibited (e.g., cross-examinations, Eades, 2007, as cited in Gardner et al. 2009), there may be very little accommodation. Differences can be treated as defiance on the part of the witness. Yet in political debates (another potentially volatile interaction but where participants have more equal status), parties adjust their volume and speech rate and have considerably more overlap (including interruptions) in the service of having their views heard instead of the opponents'. Whilst it is beyond the scope of this paper to consider these theories in detail or to conclude why or at what point interlocutors adjust their styles, this field of linguistic knowledge provides further evidence for the treatment of silences as contextual rather than directly cultural. Cross-cultural studies on silence in similar interactional situations could further clarify the roles that cultural and situational contexts play in the presence, use, and treatment of silence in human interaction.

## ***2.2. Orientation to the larger organisations of turns and sequences***

Much of the silence in this data is unmarked and unproblematic. By this, we mean that participants do not treat the silence as signalling a forthcoming dispreferred response, as non-uptake, as a possible hearing problem, or as a possible understanding problem. In the case of the former two, the first speaker may produce increments or rephrase the first pair part in the service of receiving a preferred response. In the latter two, the first speaker may initiate self-repair. In all of these instances, the silence is an accountable action, whether the second speaker's silence is due to actively doing being silent or to the first speaker's difficulty communicating. Extracts 7 and 8 show examples of marked or problematic silences.

- (7)  
 1 Paula: .hhhooh  
 2 Leif: =Should I smile a lot  
 3 >> (0.5)  
 4 Paula: Pardon  
 5 >> (0.5)  
 6 Paula: Sure

- (8)  
 1 Paula: You're still the dinner (0.6) cooker aren't you  
 2 >> (1.1)  
 3 Paula: (I mean) at your house?  
 4 >> (2.1)  
 5 Paula: >> Maybe not on (0.5) on uh (0.2) >Thursday nights



6 >> (0.9)  
 7 Leif: On Monday and Thursday nights there seems to be a  
 8 reproachment

With unmarked and unproblematic silence, on the other hand, participants show no orientation to such issues but rather continue the expected actions. The differences in what is treated as unmarked and unproblematic silence between the canonical data and that which I (as well as Gardner et al.) present could lead one to the conclusion that the interlocutors are not orienting to turn-taking or sequence organisation. However, the presence of silence and the resumption of talk are strongly organised around both turn-taking and sequence organisation:

A. Silences before the possible completion of a TCU (pauses) tend to result in the same speaker continuing or restarting the TCU that was in progress before the pause.

(9)  
 1 Leif: S↓o:: instead of- (0.5) calling it by its full name  
 2 I just called it recrui:t (0.6) do:t (1.0) do::c

(10)  
 1 Leif: [Well either either or because it was coming (1.8)  
 2 that way it probably (0.7) it probably didn't (1.9)

(11)  
 1 Paula: hh now remember that (0.4) ih- I don't have a TV  
 2 thing tuh (0.4) show us through it (0.6) >to look at

In all three of these extracts, silences occur at points in which the TCU has not been brought to possible completion, and the current speaker continues following the silence. In extracts 10 (line 2) and 11 (line 2), we also have silence before first position self-initiated repair, 'it probably (0.7) it probably didn't' and 'show us through it (0.6) >to look at'.

This pattern shows orientation to the points at which turns come to possible completion as primarily based on grammar and syntax and that speakers have a right to bring turns to possible completion before another speaker takes the floor.

B. Silences after the possible completion of a TCU to which a responsive action is relevant but before a sequence has been closed (gaps) tend to result in the prior speaker producing an increment or the recipient producing a sequentially relevant response.

(12)  
 1 Paula: Otherwise I have to get a little (1.1) TV °n I  
 2 don't really have one°  
 3 (0.6)  
 4 Leif: Television  
 5 >> (0.5)  
 6 Paula: Yeah  
 7 >> (1.3)  
 8 Leif: £That's okay£

In extract 12, there are two gaps at points in which there is a relevant response. Line 5, following Leif's other-initiated repair, shows a gap of 0.5 second that is ended by Paula responding affirmatively (line 6). Following Paula's confirmation, there is a 1.3 second gap that is ended by Leif's 'that's okay'. This thus closes the sequence that was begun by Paula regarding the availability of a television for viewing the tapes that they are recording.

(13)

- 1 Paula: You're still the dinner (0.6) cooker aren't you  
 2 (1.1)  
 3 Paula:>> (I mean) at your house?  
 4 (2.1)  
 5 Paula:>> Maybe not on (0.5) on uh (0.2) >Thursday nights  
 6 (0.9)  
 7 Leif: On Monday and Thursday nights there seems to be a  
 8 reproachment

Extract 13 is the same as extract 8 that we have previously seen. In this extract, Paula is enquiring about Leif's cooking responsibilities to his family. Line 1, the original enquiry, does not receive a response from Leif before Paula speaks again (line 3), adding detail to line 1. Then after 2.1 seconds, she produces another turn, giving an alternative at line 5. This finally receives a relevant response from Leif at line 7.

This pattern, when compared to treatment of pauses, shows orientation to the possible completion of a TCU in which the initiating action is done. It also demonstrates orientation to the sequential relevance of initiating and responsive actions, including repair insert sequences.

C. Silences after the possible completion of a TCU and after a sequence has been closed (lapses) may result in either speaker beginning a new sequence.

(14)

- 1 Paula: Zoning out in front of the TV  
 2 Leif: Habitual sex uh (.) I don't know  
 3 >> (0.9)  
 4 Leif: >> hh but [uh  
 5 Paula: [((clears throat)) scuse me  
 6 >> (3.0)  
 7 Leif: >> I I I've been thinking um hmmneh a lot about .hh  
 8 that last month or so I've also been having a lot  
 9 of sleep in my eyes the last couple three weeks

In extract 14, the participants have been talking about alternatives to drinking in a humorous tone. Lines 1 and 2 are the end of this topic, and line 3 is a 0.9 second lapse. Leif ends this lapse apparently to begin a new topic that is cut off by Paula having to clear her throat (lines 4-5). After Paula has cleared her throat, there is a 3 second lapse (line 6), and Leif moves to begin a new topic at line 7.

- (15)
- 1 Paula: [What about it was so seductive  
 2 (0.9)  
 3 Leif: .hhhh WI I've been thinking about that,  
 4 (0.8)  
 5 Leif: ↑I really have  
 6 (1.6)  
 7 Leif: Um  
 8 (1.8)  
 9 Leif: MOSTly what I didn- (3.3) mostly what I don't miss  
 10 (2.1) is disappointing myself  
 11 >> (0.7)  
 12 Paula: >> Hm

Extract 15 follows from extract 3. In the current extract, Paula has asked (line 1) what was so seductive about drinking. In lines 2-10, Leif formulates and states his answer (and is seen to be doing thinking during the silences). After Leif's answer, there is a 0.7 second lapse (line 11), which Paula ends with 'hm' at line 12. The topic is then, several seconds and another 'hm' later, picked up again by Leif (data not shown).

This pattern shows orientation to sequence closure as well as the points at which a new sequence may begin. Whether there is a lapse or a new sequence is beyond the scope of this report. Also, the presence of lapses is certainly more common in co-present data in which the main activity is attention to a stimulus or physical action rather than talking. Note that in this corpus of psychotherapy data, along with a tendency for long gaps and pauses, lapses tend to be of similar length or even shorter than gaps and pauses.

Rather than not orienting to turn-taking and sequence organisation or having an entirely different systematic, they are orienting so strongly that intra-turn silences (A) are treated as pauses rather than abandoned TCUs; inter-turn silences where a relevant response should be produced (B) end in pursuit, second pair part, or repair; and inter-turn silences where a sequence has closed (C) are treated as appropriate places for either party to begin a new sequence. This is strong evidence in support of the robustness of the organisation of turn allocation and turn design (Sacks, Schegloff, & Jefferson 1974) as well as the organisation of sequences (Schegloff 2007).

A detailed analysis of the meaning of silences is beyond the scope of this report. Certainly there are cases in which silences are in the service of 'doing not responding' or 'doing thinking' (as in Extract 15). However, I have chosen to focus on the unproblematic nature of silences that would in many data sets indicate problems or markedness in their very length. I now move into a discussion of the difficulties and possible solutions for transcribing such interactions without giving the appearance of interactional troubles and without sacrificing precision or accuracy.

### 3. Considerations for transcription

Video data enables analysts to have information about what is actually happening during silences of any length as well as during speech or other sounds. The non-verbal actions present in audio data may be overlooked (as in the case of gestures before

speaking, which could be wrongly interpreted as a delayed response) or inferred (as in the case of giving an item or opening a window). In either of these cases, researchers working with only audio data may not have a complete picture of the interaction. In working with video data, different styles of transcription and level of detail may be appropriate for non-vocal activities depending on the focus of the research. There are currently many variations on how gesture and non-vocal behaviour are transcribed, most of which are variations on Goodwin's (1979) basic approach (e.g., Streeck 2002; Stribling, Rae, & Dickerson 2007) and may be supplemented with pictorial representations or photographs (e.g., Goodwin 2002). The wide variety of research that involves non-vocal behaviour necessitates transcription styles suited particularly for the research at hand. For this reason, I've chosen to focus on the transcription of silence itself, which can be done in combination with an appropriate transcription style for the non-vocal behaviour that is relevant to the analysis.

One difficulty in transcribing data with slower than canonical speech or longer than canonical silences is that what is in effect a barely noticeable silence in context can be quite long when compared with canonical data. Analysts who are familiar with canonical data may interpret such transcribed silences as doing things they are not. For instance, quite a short silence (a micropause, canonically less than 0.1 second) in some of the data presented is roughly 0.4 second and a typical pause or gap lies in the range of 0.7 to 2.0 seconds. Whilst the participants do not treat these as long silences, analysts who are accustomed to working with more canonical data may, particularly when working with only transcripts, as one often does when reading a journal article or book chapter. Psathas and Anderson (1990) state that the relative length compared to the syllable length and silence between words (components of speech rate) is more important than the actual length of a spate of silence. There are a number of options that a transcriptionist can use to convey the data accurately in time and meaning, although a balance must be struck according to the focus of the research. I will now outline some of the options available and discuss the benefits and drawbacks of each.

### **3.1. *Giving approximate values to silences by using a fixed rating system***

This approach can be subdivided into three methods.

#### **A. Defined levels**

e.g., less than 0.5 sec, 0.5-1.5 sec, and above 1.5 sec represented as (-), (- -), and (- - -)  
This method allows for data with fairly discrete differences but with unequal intervals to be characterised as short, medium, and long silences (or as many levels as appropriate).

#### **B. Defined levels with specification**

e.g., less than 0.5 sec, 0.5-1.5 sec, and specifying lengths over 1.5 sec represented as (-), (- -), and (4.7)

This method allows for data with fairly discrete differences but with potentially very long silences to be represented compactly.

#### **C. Increments of 0.5 sec**

e.g., less than 0.5 sec, 0.5-1.0 sec, 1.0-1.5 sec, etc. represented as (-), (- -), (- - -), etc.

This method is similar to the 0.1 sec based system that is widely used in conversation analytic transcripts when symbols are used instead of numeric values to represent lengths of silences. However, it allows for a less precise interval. For very long silences,

a + can be used at 5 seconds, much the way it is used at 1 second in the traditional 0.1 sec based method.

These methods attempt to capture the approximate length of a silence relative to itself and not the surrounding talk. They simplify the usual CA transcription of silences by making the interval of interest larger and therefore less precise. However, care must be taken to develop intervals that are meaningful in light of the data. There may be quite a difference in how a silence at the beginning of an interval and a silence at the end of an interval are treated by the participants. These methods can also be more difficult to understand, when the transcriptionist uses symbols that are widely known to mean other measurements. I have used dashes here as a matter of convenience and aesthetic. As long as the transcription conventions are made clear in a given manuscript, one simply needs to consider the adjustment that readers must make to different meanings or new symbols.

### ***3.2. Using a conversation-based metric for timing silences***

This method was pioneered by Wilson & Wilson (2005) who, drawing on Wilson & Zimmerman (1986)'s experiments on the regular rate that varies between conversations with which turn-taking options are recycled in the SSJ model, argue that neuro-oscillations in each of the participants that are kept in synchrony via the rate of syllable production are responsible for such regularity and indeed is necessary for coherent conversation. They argue that when participants speak slowly, i.e. the rate of syllable production is low, both silences and sound production are stretched and that there should not be long silences in moderate or quick speech.

Wilson & Wilson (2005) go on to suggest that silence should be transcribed such that each beat of silence is relative to the surrounding talk. For example, (0.7) would represent 7 relative beats of silence and a micropause, as the participants treat it but regardless of length, as (.), with of course a note about the alteration to the transcription convention. At present there is no precise metric by which to make these judgements, but it is in development.

This paper does not argue for the use of a speech rate metric for timing silences at present. Although this may be appropriate for some conversations and participants, there are challenges that this corpus presents to such an approach:

- A. Participants' speech production rate varies across the data within a conversation and even within turns.
- B. Lengthy silences are present regardless of the rate of syllable production and are not treated as problematic. Although they may be lessened in some contexts, these shorter or less frequent silences are still not in line with canonical examples.
- C. This theory does not account for coherent conversations between parties whose syllable production rates differ (although note that people tend to adjust their speaking style to more match their interlocutors, as mentioned in Section 2.1).

### ***3.3. Taking the mean silence between words during smoothly flowing talk and subtracting this from timed silences.***

This entails timing the silences as one would normally when doing conversation analytic transcription as well as the silences between utterances during smooth spates of talk. The mean of the silences during smooth spates of talk is then subtracted from every recorded silence to give what is expected to be the real ‘silence’ as recognised by the participants. This can be a very tedious process and is not necessarily an accurate depiction despite being precise.

### ***3.4. Timing silences as they are and conveying in a description of the material that longer than canonical silences are typical of the interaction.***

This method retains the standard format of CA transcripts and is therefore easy for an analyst to read without keeping in mind new symbols or new uses of existing symbols. It does not convey in the transcript the nature of the silences in relation to other talk or their typicality, but instead asks the reader to bear in mind that lengthy silences are nothing out of the ordinary and do not necessarily mark a forthcoming dispreferred response.

## **4. Conclusions**

I have shown that lengthy silences can be present in conversation as an unproblematic and regularly occurring feature. I have also shown that despite its regularity, it is subject to local interactional contingencies and may be present in the same environment as precision speaker transition and even overlap. I have focused on a part of the corpus that has particularly long silences in order to compare situational and structural differences between the same participants. Rather than generalising the results of this study to all American English or all therapeutic situations, the objective of this paper has been to demonstrate that although early conversation analytic research as well as Jefferson’s 1988 proposed metric was focused on American English, there are contexts in which the standard maximum silence of one second does not hold. This is in line with findings in Australian English and Garrwa. Moreover, even though lengthy silences are present, there is a strong orientation to turn allocation and sequence organisation.

Taking into account that some comparatively long silences may not mean the same as silences of similar length in many other English transcripts, I have addressed the problem of transcribing such silences both accurately and meaningfully and offered some solutions that readers may adapt to their own data. I have argued for or against some approaches more strongly than others. However, by no means should these solutions be taken as better than any other or better than each other overall. Each has its strengths and weaknesses and is more appropriate for some data than others. Readers may even find that different styles prove better for showing different phenomena.

## Acknowledgements

I would like to thank the participants who chose to make their private interactions available for study, and my supervisor, John Rae, for reading early drafts and making valuable suggestions along with Rowena Viney. Also, thanks are due to the users of the Language Use listserv for engaging in discussions both on the list and privately on the transcription of silences and suggesting readings. This paper is a direct result of these most interesting discussions and encouragement.

## References

- Bateson, G. (1935) Culture contact and schismogenesis. *Man* 35: 178-183.
- Dickerson, P., P. Stribling, and J. Rae (2007) Tapping into interaction: How children with autistic spectrum disorders design and place tapping in relation to activities in progress. *Gesture* 7.3: 271-303.
- Gardner, R., R. Fitzgerald, and I. Mushin (2009) The underlying orderliness in turn-taking: Examples from Australian talk. *Australian Journal of Communication* 36.3: 65-90.
- Giles, H., and J.M. Wiemann (1987) Language, social comparison and power. In C.R. Berger and S.H. Chaffee (eds.), *The Handbook of Communication Science*. Newbury Park, CA: Sage, pp. 350-384.
- Goodwin, C. (1979) The interactive construction of a sentence in natural conversation. In G. Psathas (ed.), *Everyday Language: Studies in Ethnomethodology*. New York: Irvington, pp. 97-121.
- Goodwin, C. (2002) Conversational frameworks for the accomplishment of meaning in aphasia. In C. Goodwin (ed.), *Conversation and Brain Damage*. Oxford: Oxford University Press, pp. 90-116.
- Gourley, L.E. (1973) Capital B for Black. *American Psychologist* 30.2: 181.
- Jefferson, G. (1988) Notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation. In D. Roger and P. Bull (eds.), *Conversation: An interdisciplinary perspective*. Clevedon, UK: Multilingual Matters, pp. 167-196.
- McWhorter, J. (2004, September 18) Call me Black. That's with a capital B. Not African-American. *Houston Chronicle*. Retrieved from <http://www.chron.com/disp/story.mpl/editorial/outlook/2802824.html>
- Pomerantz, A.M. (1984) Agreeing and disagreeing with assessment: Some features of preferred/dis-preferred turn shapes. In J.M. Atkinson and J. Heritage (eds.), *Structure of Social Action: Studies in Conversation Analysis*. Cambridge: Cambridge University Press, pp. 57-101.
- Prince, R. (1997) Is talk cheap or silence golden: The negotiation of activity in psychotherapy. *Psychoanalytic Inquiry* 17: 549-558.
- Psathas, G., and T. Anderson (1990) The 'practices' of transcription in conversation analysis. *Semiotica* 78: 75-99.
- Sacks, H., E.A. Schegloff, and G. Jefferson (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50: 696-735.
- Schegloff, E.A. (2007) *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.

Streeck, J. (2002) Grammars, words, and embodied meanings: On the uses and evolution of so and like. *Journal of Communication* 52: 581–596.

Street, R., and H. Giles (1982) Speech accommodation theory: A social cognitive approach to language and speech behavior. In M. Roloff & C.R. Berger (eds.), *Social Cognition and Communication*. Beverly Hills, CA: Sage, pp. 193-226.

Stribling, P., J. Rae, and P. Dickerson (2007) Two forms of spoken repetition in a girl with autism. *International Journal of Language & Communication Disorders* 42: 427–444.

Tannen, D. (1984) The pragmatics of cross-cultural communication. *Applied Linguistics* 5: 189-195.

This Wicked Day (2010, June 16) Capitalisation is a political act. [Web log comment] Retrieved from <http://wickedday.wordpress.com/2010/06/16/capitalisation-is-a-political-act/> (2011, February 12)

Wilson, M., and T.P. Wilson (2005) An oscillator model of the timing of turn-taking. *Psychonomic Bulletin & Review* 12: 957-968.

Wilson, T.P., and D.H. Zimmerman (1986) The structure of silence between turns in two-party conversation. *Discourse Processes* 9: 375-390.