

Assessing fluency: Are the criteria fair?

*¹Indika Liyanage and Rod Gardner

Griffith Institute for Educational Research, Griffith University, Queensland, Australia

Abstract

Fluency is an important aspect of verbal interaction and is assessed formally using a range of language features. However, there is certain bias associated with how these features are used to measure fluency or disfluency with native (L1) and non-native (L2) speakers of a language. With L1, language features such as pausing, silences and self-repairs, which are usually indicative of disfluency in the interactions of L2, are regarded as compensating tools for interactional problems, and are thus considered an important interactional resource. In this paper, we argue that in L2 talk these features are not always indicative of disfluencies, but are often rather the results of good interactional practice. We use close transcription and Conversation Analysis tools to focus on pausing, silences and self-repairs in the speech samples in two data sets: the first set comprise the interactions of two L1 speakers of English who were teaching English to a group of Korean students, and the second set comprised the mock IELTS interviews of I-Kiribati² students who were preparing for subsequent tertiary studies overseas. We found that the use of pausing, silences and self-repairs, as well as the distribution and frequency of them, were often very similar in the two data sets suggesting that a superficial observance of disfluency markers in the case of L2 speakers hides some communicative and interactional practices that may in fact display a high level of competence. We discuss these findings highlighting the caution that needs to be exercised when making fluency judgements.

Keywords: Fluency; disfluency; proficiency testing, measures of fluency

¹ Corresponding author. Email: iliyanage@griffith.edu.au

² Kiribati comprises a landmass of just over eight hundred kilometers spreading across 33 atoll islands which lie from east of Nauru to French Polynesia and is situated in the Central Pacific Ocean. The country gained independence from the British in 1979. I-Kiribati – the people of Kiribati.

Introduction

Over thirty years ago, Filmore (1979) presented a broad view of fluency, identifying four kinds, namely, “the ability to talk at length with few pauses,... the ability to talk in coherent, reasoned, and ‘semantically dense’ sentences, ... the ability to have appropriate things to say in a wide range of contexts, ... [and] the ability some people have to be creative and imaginative in their language use” (p. 93). Most of the psycholinguistic research into fluency however has focussed only on the first of these, and our critique of this work in this article is in part a reflection of the narrowness of this focus. Notwithstanding this, researchers within psycholinguistics have different views on what constitutes oral fluency. For example, some see it as a complex interaction between listening and speaking (Derwing, Munro, & Thomson, 2008; Derwing, Rossiter, Munro, & Thomson, 2004) mediated by activity type and context, while others (Chambers, 1997) consider it to be a production skill reflecting overall language proficiency. Fluency is also considered a speaker attribute. For example, Lennon (2000) sees fluency as the general command of the spoken language demonstrated in “rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal conditions of on-line processing” (p.26). Fluency is also considered a perception on the part of the listener. For example, Lennon (1990) takes a view that fluency can be seen as “an impression on the listener’s part that the psycholinguistic processes of speech planning and production are functioning easily and efficiently” (p.391).

Despite these differing views, there is also widespread agreement in the literature that fluency is smooth, accurate and automatic production of language. It has been characterised as an “automatic procedural skill” (Schmidt, 1992, p. 358) or as automaticity of production (Derwing et al., 2008; Derwing et al., 2004; Towell, Hawkins, & Bazergui, 1996) resulting in “smooth and rapid production of utterances, without undue hesitations and pauses, that emerges from constant use and repetitive practice” (Gatbonton & Segalowitz, 2005, p. 326).

Kormos and Denés (2004, p. 161) suggest fluency “is best conceived of as fast, smooth and accurate performance.”

These views mean that fluency extends to competencies broader than mere linguistic knowledge, and the term has very widespread use (Chambers, 1997). What constitutes fluency and its measurement have been and continues to be a matter of debate (Chambers, 1997; Kormos & Dénes, 2004; Riggensbach, 1991). Kormos and Dénes (2004, pp.148-149) summarise four foci in the studies of language features that contribute to perceptions of fluency: temporal aspects, temporal aspects linked to interactive features, phonological aspects, and analysis of formulaic speech.

Research continues to pursue clearer definitions of fluency and features that lead to perceptions of fluency. These features can be applied to both L1 and L2 speakers. However, perceptions of fluency or its measurement relate most generally only to L2 speakers (Riggensbach, 1991) in a rather unjust and complicated manner. L2 speakers are the receivers of constant formal and informal fluency judgements. For example, as Derwing, et al. (2004, p. 676) point out, in informal settings “taxi drivers, shopkeepers, and telephone receptionists , among others, make judgements daily and with ease regarding L2 fluency.” In formal settings, like proficiency testing, L2 speakers’ oral proficiency is judged against criteria such as pauses, silences and hesitations, which in the case of L1 speakers are not markers of disfluency but normal solutions to problems such as managing turn taking, looking for an appropriate word or organising content. Sometimes, further criteria are used in measuring L2 fluency. For example, disfluencies in L1 speech are associated with the non-alignment of intention and meaning, but in L2 production, this also extends to struggling with morphological and syntactic realisations (Chambers, 1997).

However, more than forty years ago, Goldman Eisler (1968) argued that pauses and hesitations are not necessarily indicative of lack of fluency in speech, but rather that they occur when new speech is being formulated and high levels of cognitive activity are required. Conversely, they do not occur around habitual speech that requires low levels of cognitive activity. She also found that pauses and hesitations tended to occur before words that are less predictable in that context. McCarthy (2010) extends the notion of fluency further by arguing that fluency is not just the production of a single speaker, but the joint, interactive production of two or more speakers. While much of the research and assessment of L2 fluency is about measurements or perceptions of speech rate and pausing, increasingly there is recognition that fluency is an interactive phenomenon. In this spirit, we argue in this paper that the use of commonly perceived disfluency markers such as pausing, silences and self-repairs by L2 speakers are in fact not always indicative of disfluency; they are rather features resulting from good interactional practices and hence they should be viewed cautiously when fluency judgements are being made for assessment purposes. In the next section, we review the literature relevant to ideas and concepts foreshadowed in this section.

Literature review

According to the relevant literature, in listeners who are L1 speakers of the TL, varying impressions of disfluency in L2 speakers can be created by the absence of features of automaticity. In general terms, disfluency is manifested in temporal intraturn features such as “filled pauses³, excessive pausing, pausing in inappropriate places, false starts and a slow speaking rate” (Derwing et al., 2008, p. 360). While a core set of temporal variables is used

³ The term ‘Filled Pause’ is problematic. Vocalisations such as ‘*uhm*’ do a variety of sophisticated interactional work that is more than merely filling what would otherwise be a silence. The literature we are citing, however, uses this term widely. We discuss below the ways in which such vocalisations are used. This will have implications for the ways in which we conceptualise fluency.

by many researchers in the analysis of speech samples, there are variations in the range of features and measurement of fluency. In addition, the sets of variables differ from study to study. For example, Kormos and Dénes (2004) analysed ten distinct temporal variables, while Towell et al (1996) used five. These temporal features are operationalised in the research literature in a range of variables open to measurement. While broad agreement exists on what constitutes fluency, there is less agreement on the specific features that contribute to perceptions of disfluency.

Researchers have considered a broad range of such features. These include amount of speech, which is the raw count of all words/semantic units per minute, and may include ‘filled pauses’ and partial words (Riggenbach, 1991); speech or speaking rate, which is some measure of the number of syllables produced in a unit of time, e.g. syllables per minute (Towell et al., 1996), or words or semantic units per minute (Kormos & Dénes, 2004; Riggenbach, 1991); articulation rate, which is the mean number of syllables or phonemes produced per second (Towell et al., 1996), and may include ‘filled pauses’ and partial words (Kormos & Dénes, 2004); a related measure is pruned syllables per second (Derwing et al., 2004) which excludes for example, self-corrections, false starts, and asides; phonation/time ratio measures the amount of time spent speaking as a proportion of the total time taken to produce an utterance (Kormos & Dénes, 2004; Towell et al., 1996); mean length of runs refers to the average number of syllables produced between pauses of a designated length (Kormos & Dénes, 2004; Towell et al., 1996); a more global measure is the number of disfluencies per minute (Kormos & Dénes, 2004), that is, the total number of repetitions, restarts, and repairs per unit of time; pace is the number of stressed words per minute, and space is a measure of stressed words as a proportion of the total number of words (Kormos & Dénes, 2004). Other studies have examined the role of prosodic features of speech (Derwing

et al., 2004) or of specific features such as pitch (Wennerstrom, 2000); self-initiated repairs include retraced and unretraced restarts (Riggenbach, 1991).

Virtually all studies include various types of pausing as a major contributing factor in perceptions of fluency. Pauses or some hesitation phenomena (Kormos & Dénes, 2004; Riggenbach, 1991) are reflected in articulation rates and phonation time ratios but are more complex in practice. Riggenbach's (1991) study distinguishes between 'unfilled and filled pauses', whereby unfilled, or silent, pauses include micropauses (0.2 seconds or less), hesitations (0.3 – 0.4 seconds) and 'unfilled pauses' (0.5 seconds or greater). 'Filled pauses' make no semantic contribution to the talk, and include 'fillers' that can be non-lexical (e.g., *um, ah*), lexical (e.g., *you know*) and sound stretches (vowel elongations greater than 0.3 seconds). Chambers (1997) considers pauses to be part of 'natural' speech in L1 speakers, but states that they need to be distinguished from hesitations that do not occur at semantic junctures or intakes of breath and which can thus be regarded as 'unnatural'. Such hesitations, including 'filled' and 'unfilled' pauses need to be judged according to the environments in which they occur. That is, in terms of pauses, one factor that distinguishes L1 from L2 speakers is the placement of them. Moreover, pausing in the speech of L2 speakers may also be due to learners transferring patterns of pausing from their L1. Chambers (1997) further discusses frequency of pausing as a factor in the perception of fluency. Riggenbach (1991) makes the point that clusters of more than three hesitations may impact on listeners' perceptions of fluency. Thus, even 'natural' pauses surrounded by hesitations can become part of clusters that are judged to be disfluent (e.g., Raupach, 1980).

Towell et al. (1996), however, suggest that perception of pausing as a factor in the measurement of fluency may be much more complex than this. They note that:

... pausing covers many factors. Change in such measurements may be caused by the demands of a particular task, it may be a characteristic of an individual, it may be a sign of individuals having difficulty knowing what to say, it may be a sign of having difficulty in knowing how to verbalize what is already in the mind, it may be to do with establishing the correct 'balance' between length of utterance and linguistic structure (p. 93).

Fluency, it has been claimed, is a feature of performance that is variable within specific contexts (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001). Learners can demonstrate variations in speech performance in response to variations in cognitive load within utterances or interactions (Raupach, 1980), e.g., variations in fluency are attributable to demands of differing tasks (Derwing et al., 2004). Studies that focus on the development of fluency often assume that beginners and low-proficiency learners are “relatively homogeneous” (Derwing et al., 2004, p. 674) in exhibiting a lack of fluency due to limited knowledge of syntactic and lexical aspects of the TL. Ranta and Derwing (2000), however, found variations in the fluency of beginners with similar levels of linguistic knowledge. Derwing, et al. (2004) may partially account for this with their finding that fluency of low-proficiency learners differed significantly depending on task type. Although the relationship is poorly understood (Oviatt, 1994), it is believed disfluencies increase in line with planning demands (Levelt, 1989) and that planning time results in greater fluency as well as grammatical complexity (Derwing et al., 2004). It is claimed that in interactive talk, sociolinguistic and affective factors, including speakers’ perceptions of listeners’ expectations, play a significant role (Kormos & Dénes, 2004; Riggenbach, 1991). On this last point, a complex set of linguistic, communicative and psychosocial situational variables modelled as the willingness of speakers to communicate (MacIntyre, Clément, Dörnyei, & Noels, 1998) offers insight into contextual factors as influences on L2 fluency.

The interplay between temporal and interactive features in context suggests that “in order for there to be fluency, then, it appears that many different conditions have to be met” (Riggenbach, 1991, p. 439) and that difficulties in one area may be enough to influence perceptions of L2 speakers as disfluent. Temporal features as conflated in the measure of pruned syllables per second, which exclude self-repairs, ‘filled’ pauses and asides, were found to be reliable indicators of perceptions by untrained raters of fluency of low-proficiency learners (Derwing et al., 2004), although others have argued that there is more to fluency than temporal measures such the absence of pauses in speech (Lennon, 2000).

Some models of fluency (Lennon, 1990) propose core and peripheral variables in the judgement of fluency. Kormos and Denés (2004) suggest that speech rate was a good indicator of fluency, as well as mean length of runs. The same study concluded that pace, or more specifically patterns of stress, was the key temporal characteristic of fluent speech. This however conflicts with the findings of Chambers (1997), who suggests pauses as a core variable, arguing that L2 speaker development towards greater fluency involves pausing less often and at ‘appropriate’ places in utterances, rather than speaking faster. Freed (1995) on the other hand found that speech rate and pausing were the most frequently used indicators in fluency judgements by the six raters in her study.

Riggenbach (1991), however, found that L2 speakers exhibiting low frequency of silences and acceptable rate of speech can still be judged as disfluent by listeners: “it may be that the restricted definition of fluent speech as simply smooth and continuous becomes applicable only beyond a certain level of linguistic proficiency” (Chambers, 1997, p. 540). Chambers (1997) concludes that temporal features alone provide a restricted picture of what constitutes fluency that can be “refined by an analysis of the language, in particular its syntactic features” (p. 543). Further evidence comes from Riggenbach’s (1991) analysis of another intra-turn feature, repair, in very fluent and very disfluent L2 speakers. She urges

caution in considering intraturn repair as a marker of lack of fluency, as she found that an L2 speaker judged globally to be highly fluent produced a high incidence of repair which had in fact resulted in “grammatically or semantically improved utterances” (1991: 433). Another L2 speaker judged to be very disfluent displayed a low incidence of repair, which may have indicated a lack of awareness of problematic structures.

Raupach (1980) reported that an increase in mean length of runs led to an increase in fluency ratings. This view is supported by findings of a study of a small group of advanced foreign language learners of French (Towell et al., 1996) that found that increases in fluency were:

not the result of a quantitative reduction in the amount of pausing that subjects do, nor in the increase in the speed with which they articulate what they say. Rather, there is an increase in the length and complexity of the linguistic units which are uttered between pauses. This suggests that what has changed is the rapidity with which syntactic and discourse knowledge can be accessed for on-line speech production (Towell et al., 1996, pp. 112-113) As noted above, some of the research on fluency acknowledges the effects of variables such as the type of task or type of spoken text (Derwing et al., 2008; Towell et al., 1996), time available for the planning of speech and the planning process (Nation, 1989), time spent in an L2 environment, speaker variables such as self-monitoring, whether the context of oral production is monologic or interactive (Hilton, 2008), status of raters as trained or not (Derwing et al., 2004; Riggensbach, 1991), and whether assessments were made face-to-face or through recordings of the speech (Derwing et al., 2004).

Among these cited studies, it can be concluded that there is a reasonable consensus that the most significant factors in judgements of L2 fluency are pausing in its various guises, rate of

speech, mean length of run, and possibly self-repair. At this point, it is worth asking how these features affect judgements of L1 fluency.

In cognitive terms, fluency in L1 is argued by some to be dependent upon automaticity of the components underlying processing and production of speech in an individual (Levelt, 1989). Identifiable disfluencies in L1 speech are commonly believed to be observable signs of problems experienced by speakers with alignment of intention and meaning (Chambers, 1997), that is, problems of planning utterances (Bortfeld et al., 2001; Clark, 2002). However, in interactive contexts automaticity of production of speakers is subject to the influence of sociolinguistic and psychosocial variables. This alternative view holds that disfluencies are solutions to problems, rather than the problems themselves (Clark, 2002, p. 13). Clark argues that the strategies adopted by speakers to synchronise perception, analysis and semantic understanding of utterances by their addressees constitute fluency and amongst these strategies are features that are conventionally regarded as disfluent.

They are genuine signs – collateral signals – that speakers design and produce with skill. Taken together, they constitute an elegant solution to a complicated problem: how speakers and addressees are to coordinate their communicative actions while carrying out the official business of their discourse (Clark, 2002, p. 13).

For example, what Clark calls as ‘conversational fillers’ (‘filled pauses’ above may be a resource for or consequence of interpersonal coordination in L1 spoken interaction (Bortfeld et al., 2001). Guillot (1999) makes a similar point about self-repetition by L2 speakers which is seen by many as a marker of disfluent speech, whereas it may be perceived by others as a way of buying time that creates a perception of fluency on the part of addressees. Similarly, some silence and hesitations are not only acceptable, but are even essential aspects of L1 speech. For example, Riegenbach (1991) notes that some of these

pauses can be judged to be natural (and fluent) when they occur at the juncture of units of meaning or enable intakes of breath.

It can be assumed that second language speakers will display varying degrees of lack of fluency, from native-like to very hesitant and disjointed. Given what we argued in the previous section, a problem we perceive is how to distinguish between what we might term ‘native-like disfluency’, and disfluency that is caused by lack (in some form) of proficiency in the language. What we can glean from the previous section is that certain apparent disfluencies have interactional motivations, and some other disfluencies result from what we might call processing problems, whilst yet others may have their source in a combination of these. In the analyses that follow, we will be examining extracts of talk in which we find examples that the fluency literature refers to as, “pausing or hesitation at inappropriate places”, “excessive or undue pausing and hesitation”, and “excessive self-repair”. Within the conversation analytic literature on spoken interaction, there is a considerable body of work on pausing, silence and self-repair.

First, it may not be an easy matter to determine what an ‘inappropriate place’ to pause is. Silences, pauses and hesitations are complex phenomena in talk. One regularly finds, for instance, that L1 speakers pause in the middle of an utterance where they can exercise control of their turn by putting their utterance on hold for a moment before it is completed, thereby ensuring that they do not lose the floor (Schegloff, 1979). Silences also regularly occur as a gap between speakers’ turns, particularly after responses such as answers or acceptances of offers (Sacks, Schegloff, & Jefferson, 1974). ‘Filled pauses’ or what we would prefer to call ‘turn-holding devices’, such as an *uhm*, are a resource that can be used at the beginning of a turn to ensure that one has the floor before one is ready to speak (Schegloff, 2010). Such examples show that hesitations and pauses can in fact be a valuable resource for managing turn taking.

Pausing and hesitations in L1 talk can cluster in certain phases of a spoken interaction, and there may be good reasons for such clustering. To decide what is 'excessive' or 'undue' can be problematic if one does not take into account where such clusters occur in natural L1 conversation. For example, when an invitation, request or complaint is being rejected, one regularly finds pauses and hesitations before or during such rejections (Schegloff, 2007). Silences are also interspersed in talk on delicate topics (Schegloff, 1990). Searching for a word that one cannot recall immediately is another place at which hesitations (turn holding devices such as 'uhm') accumulate (Brouwer, 2003). Longer silences are found when people are together for long periods of time or when they are involved in an activity such as cooking (Goodwin, 1981; Mushin & Gardner, 2008; Schegloff, 2010), or when a topic in a conversation is running down (topic attrition) (Jefferson, 1984). It is found that answers to 'why' or 'how' questions are often preceded by silences and include turn holding devices in L1 talk (Schegloff, 1990). Significantly, 'why' and 'how' questions are often found in L2 learner oral tests.

Occurrence of self-repair in L1 talk is also a complex phenomenon, and has been shown to be more than merely a result of cognitive processing problems. For example, a speaker may self-repair because of a perception that the way the utterance is being formed is inappropriate for the addressee (Schegloff, 1979; Schegloff, Jefferson, & Sacks, 1977). Repetitions, too have been shown to occur systematically in certain position in spoken interaction (Schegloff, 1984, 1996).

This body of work demonstrates some of the complexity of the occurrence and clustering in L1 talk of certain features of speech considered by many researchers into L2 fluency to be markers of dysfluency. It also needs to be noted that dysfluency is, to some degree at least, a co-construction between current speaker and recipient. For example, if a listener withholds minimal responses (e.g. 'yeah', 'mm hm' 'oh' or 'that's great'), this can

disrupt the fluency of the main speaker (Gardner, 2001). Again, in many L2 oral tests, interviewers are instructed to withhold just such minimal responses.

We argue that in the interests of fairness to learners or test candidates, we need to at least consider ways of distinguishing 'natural', 'normal' or L1 types of dysfluencies from ones related to lack of proficiency in a language, even if it turns out that this is very difficult, or may even be impossible in situations such as oral tests.

We turn now to some extracts of the classroom talk of some ESL/EFL teachers whose L1 is English and some Kiribati nurses learning English in preparation for study at an Australian university. The first data set is five ESL classes with Korean L1 learners and three English L1 teachers totalling about four and a half hours. The second set of extracts is from a mock interview test with ten I-Kiribati trainee nurses preparing for IELTS and subsequently for tertiary studies in Australia. Each of these interviews is about eight minutes long. In the following section, we will focus mainly on pausing and hesitation, and self-repair. All have been closely transcribed using Jeffersonian conventions which have been simplified in the extracts presented below.

Data analysis

In this section, we present some extracts of classroom talk by teachers and students. For this study, we chose randomly ten out of 30 I-Kiribati students we had recorded. We transcribed the oral tests for each of these students in full using Jeffersonian transcription conventions. We then comprehensively analysed the data to identify all potential markers of disfluency. From these, we selected a range of examples that were typical and suited to illustrate the points raised in this paper. We demonstrate how the talk of English L1 teachers can incorporate a range of disfluency features that might be considered excessive and even inappropriate. It is not our purpose to claim that such features are found routinely in the talk

of L1 speakers, although it was not difficult to find 14 examples of very disfluent English L1 teacher talk in our data. Neither can we explain why the talk of these teachers at these points in the lesson became so disfluent, whilst in other parts of the lessons their talk was much more fluent. We might conjecture that they were breaking up their utterances into manageable chunks, or it might be that they were attempting to find simple vocabulary in simple grammatical structures for their learners, rather than more complex language that might be more natural in L1 talk. Be that as it may, contextual variables are likely to have played their part in making this talk so disfluent.

Excessive or inappropriate pausing, hesitation and self-repair

In the first extract, the class is practising the grammatical forms ‘looks like’ and ‘looks as if’, as in ‘You look like you have seen a ghost’ or ‘It looks as if you have seen a ghost’. Students are constructing example sentences from a stimulus phrase. In this following extract, the teacher, Barry⁴, a Canadian, is attempting to give an example. However, he confuses the predictive or evidential use of ‘looks like’ (the one being practised by the students) with the more literal comparative ‘x looks like y’, as he seeks to provide an example of the form for the class. This confusion appears to contribute to the massive disfluency of his talk.

(1) K-TR14:224⁵

1 T-Barry: No I meant about Canada:. so for

⁴ All names have been changed.

⁵ These transcriptions follow Jefferson’s widely used system (e.g. Jefferson, 1984). Speakers are on the left. Underlining means that the word or syllable containing the underline is stressed. Colons show sounds that are stretched or drawled. Full stops indicate falling terminal intonation, question marks strong rises, and commas more or less level. Numbers in brackets are silences in seconds and tenths of seconds. [.hh] indicates an audible inbreath Angle brackets pointing outwards (<word>) show slowly articulated speech.

2 exam:ple·hhh Can:ada:? (1.0) uh
 3 Canada:- looks a bit li:ke hh (1.2)
 4 wint- (0.1) u:h Can:ada: (0.8) the
 5 Cana:dian weath:er, (0.2) u:h looks a bit
 6 like (0.8) the arctic.
 7 (3.4)
 8 T-Barry: Canadian wea:ther: (1.5) uh Canadian
 9 weather: l:ooks:- (0.6) like (0.3)
 10 the arctic.
 11 (0.2)
 12 T-Barry: er fee:ls li:ke- (0.3) the arctic.

Barry's talk includes eleven silent pauses, four of which are longer than one second. No fewer than nine of these pauses come at grammatically and semantically incomplete points. There are also five turn holders in lines 2, 4, 5, 8 and 9, as well as five self-repairs in lines 2, 4 (two instances), 8 and 12 in this turn. This all amounts to what can be seen as a massive accumulation of disfluency features. This has come about as he has had trouble in formulating an example that includes the grammatical form 'looks like'. However, it needs to be noted that in ordinary conversations, L1 speakers regularly pause in the middle of utterances to ensure that they keep the floor. Such pauses can be filled or unfilled (i.e. silences or turn-holders such as 'uhm'). Pausing at grammatical and semantic completion points puts a speaker in danger of losing the floor, as these are points at which other speakers may legitimately begin to speak (Sacks et al., 1974). In addition, an 'uhm' at the beginning of a turn is a way of letting others know that one has taken a turn, but at this point is not quite ready to begin speaking. It can also be noted that he uses stress very extensively, which

would affect perceptions of ‘pace’, or the number of stressed syllables per minute. However, more appears to be going on here than just turn-taking behaviour. The teacher uses all of these turn-holding resources to show he has not finished his turn, but also because he is having a great deal of trouble putting his turn together.

Our second example is from a conversational warm-up activity at the beginning of a lesson led by an American teacher of an EFL class in Korea. In the talk leading up to this excerpt, the student, Ann, a member of a church choir, was talking about what she has to do at the church. Ken picks up on her use of ‘have to’ as an incorrect verb form, but Ann challenges this by saying it is her obligation to the church, which suggests that she understands that the semantics of ‘have to’ includes the notion of obligation. In the extract presented below he appears to be looking for an appropriate alternative. But, he struggles.

(2) K-TR8:57

1 T-Ken: Okay because you:: (0.4) you pr:o:mised
2 yourself, you pro:mi:se[d,
3 S1-Ann: [Yeah
4 T-Ken: that if- (.) oh:ka:y
5 S1-Ann: Mh:m,
6 T-Ken: an obligation. (.) [you you-] promise
7 S1-Ann: [M m ,]
8 yourself. ·h s:o you: vol:untee:red.

Ken’s talk in this extract contains a pause of 0.4 seconds clause-medially in line 1, and an accumulation of self-repairs and repetitions. He repeats ‘you’ in lines 1 and 6. There are self-repairs in lines 2 (‘you promised yourself’ to ‘you promised’) and a false start in line

4 ('that if'), after which he restarts by saying 'okay' and then re-establishing his topic, 'an obligation'. He ends up by resolving his word search with 'volunteered'. In this brief extract we have a pause at a non-grammatical or semantic juncture, and a word search with two self-repairs. Ken's disfluency can be attributed to his search for an alternative to 'obligation', which turns out to be 'volunteer', but this is also possibly aggravated by Ann's challenge to the teacher when he said that she had no obligation to the church. Here we have an example which shows some of the contextual factors that need to be considered when attempting to account for breakdowns in fluency.

When we turn our attention to the Kiribati student nurse learners of English, we can examine the occurrence, nature and placement of pausing, hesitation and self-repair. The first example we present is Marta responding to the question, "Where would you like to take a holiday, and why?"

(3) A008:Marta:49:Holiday

1 Ann: <Where would you like to take a ho:liday,>
2 (0.5) [and why:..]
3 Mar: [So:rry,]
4 (0.4)
5 Ann: <Where would you like- (.) to take a
6 ho:liday; (0.7) and why:..>
7 Mar: ·HHH A:hh we:ll. ·hhh I: rilly like to take a
8 holiday? (0.8) 'e:re in Kiribati u:hm
9 originally: a:t my home i:sland?
10 (0.6)
11 Mar: uhm Maraki?

12 (0.3)

13 Mar: one of thu:h (0.2) south i:slands, here in
 14 Kiribati? and I want to go there becau:se?
 15 ·hhh (1.7) this is where my:: (0.6) mum come
 16 from.

17 (0.6)

18 Ann: Mm:.

19 (0.6)

20 Mar: he:r home (.) island, and that's why: I want
 21 to go there to (0.4) have a chance to meet
 22 her relatives the:re,
 23 (1.0)

24 Ann: Good.

25 (0.7)

The interviewer articulates the questions slowly and clearly, pausing for half a second before completing with ‘and why’. The intonation after ‘holiday’ rises slightly terminally, indicating non-completion of the question. After the 0.5 second silence, Marta says, “sorry” simultaneous with the interviewer’s “and why”. Marta has produced a repair initiation that solicits a repetition which, in the eyes of an examiner might be interpreted as a sign of disfluency. This could also suggest a lack of understanding on the part of the student. However, because of the slightly rising intonation after “holiday”, an alternative interpretation could be that she has heard an incomplete question followed by a significant pause. It is therefore plausible that Marta is demonstrating interactional competence rather than disfluency by recognising potential trouble in the delivery of the question.

Ann repeats the question in full, and Marta launches her answer relatively fluently, with only one unfilled silence of 0.8 seconds, one hesitation marker and no self-repairs. There may nevertheless be some perception of disfluency here despite the low number of disfluency markers, and one has to ask whether the relatively slow pace of delivery and grammatical errors could contribute to such perceptions. In examples 1 and 2, the teachers Barry and Ken also delivered their turns at a slow pace. During the rest of her response, Marta displays relatively few of the markers of disfluency we are focussing on; indeed in many ways her talk is more fluent than Barry's or Ken's in the examples above. There is only one turn holding token in her talk and only one silence of more than a second. There are no repetitions, no repairs and, incidentally, her length of runs (number of words produced between pauses) is high, averaging about eight words per run. Some silences occur in the middle of utterances, which as we saw in Barry's talk in extract 1, are used in talk as a resource for keeping one's turn. This again could be taken to indicate interactional competence on Marta's part rather than disfluency.

It is also worth noting that she is responding to a complex question that requires multiple utterances. She needs to decide spontaneously not only on a holiday destination, but also to think up reasons why she would want to do what she normally would not do (holidaying is not an activity that ordinary i-Kiribati normally undertake). Given these circumstances, it would be unsurprising for multiple pauses to occur as she constructs her reasons. It is perhaps more remarkable how few silences and other markers of disfluency occur in her answer.

The next extract shows a response to a question that requires an extended description of the candidate's family home. Keira is one of the i-Kiribati students whose talk includes above average incidence of the disfluency markers that we are examining.

(4) A003:Keira:55:Home

1 Ann: Can you describe your family home.
2 (2.7)
3 Kei: <De::re a::re> (0.6) three of us:.
4 (0.5)
5 Kei: Myse::lf? ·hhh an' I 'a:d u:h (.) two
6 bro:thers?
7 (1.6)
8 Kei: ex:cluding our pa:rents, so: (0.7) there's:
9 five of us in the fa:mily? ·hhh an::d (1.8)
10 we::: (3.7) we love each o:ther? (fi:rst)?
11 ·hhh bu:t now u:hm (0.6) our home is: more
12 like e:h:m::: (1.2) <more like u::hm::> (1.6)
13 a:::h (0.5) ex:tended f:amily?
14 (0.5)
15 Kei: (who-) (.) because we live in an extended
16 family, most of o:ur (0.7) most of our
17 r:e:latives living with us?
18 (1.0)
19 Kei: <Bu:t we::> (1.7) <we: enjo:::y:> doing lot of
20 thing:s?
21 (1.0)
22 Kei: so::: (1.0) <like e::hm::, (0.8) doing
23 mo::vie:s,> together, watching movies

24 toge:ther,
 25 (1.1)
 26 Kei: u::hm: (1.0) playing ga:mes:
 27 (0.4)
 28 Kei: so:: (0.5) our family home is: mo:re like
 29 a:h (0.2) happy ho:me I can say.
 30 (0.4)
 31 Ann: °Mm:°
 32 (0.6)
 33 Ann: °Good°

There are significantly more long pauses than in the previous example. First, there is a gap of 2.7 seconds before she launches her answer, which in itself can be seen as possibly time she needs to plan and construct an extended and complex answer. In the rest of her answer, she produces another five silences of one second or more, most of which occur at grammatically and semantically incomplete points. The two longest of these, at 1.8 and 3.7 seconds, occur within the same utterance, and appear to be associated with a search for something to say next. We could compare Keira's search for her next words with the teacher Ken's search for the word 'volunteer' above. There are also four long pauses of a second or more at points of grammatical and semantic completion. Note that the interviewer was instructed to provide no feedback to the student during her answer, in for example the form of minimal responses or evaluative remarks. Such instructions are of course given to many oral test interviewers in widely used proficiency tests around the world. Minimal responses are an integral part of spoken interaction, and the lack of them can be disconcerting for a speaker delivering an extended turn at talk, such as here. It is precisely at points such as where these

four silences occur that minimal responses would routinely be placed in L1 talk (Gardner, 2001). One could thus argue that these silences belong to the interviewer rather than the candidate, but is it likely that these silences will be seen as candidate's silences, and thus add to impressions of disfluency.

In this section, we have examined some examples of talk taken from TESOL classrooms and oral test situations, focussing mainly on pausing and repair. We have seen that L1 teacher talk can be strewn with such disfluency markers. These do not necessarily indicate a breakdown in fluency, though they may. Many are used to manage turn-taking or for word searches, both of which are legitimate and natural practices in spoken interaction. One can thus conclude that there are good reasons for using them. When we looked at the second language data, we also found hesitations to occur frequently. It was noted that some of these showed interactional competence at work, but in addition, it was claimed that some of the others could be attributed to the interviewer and the manner in which the interview was conducted. We can conclude that it is imperative that we consider interactive and contextual factors when making judgements about the fluency of second language speakers.

Conclusion

In this paper, we have focussed mainly on three markers of disfluency that are mentioned very widely in the research literature on fluency: silences of various kinds, filled pauses (turn holding devices), and self-repairs. Most of the literature identifies frequency of occurrence of these features as contributing to disfluency. There is little discussion of specific interactional or contextual grounds for their occurrence. We judge L2 speakers often without differentiating between functionally and communicatively effective and appropriate fluency (or disfluency) features, and tend to ignore fluency practices that are regularly found in L1

speech or talk. Nor do we usually take into account that these measures of fluency are taken during what is for most people the stressful situation of an oral interview or test.

We are not suggesting that fluency is not an issue or real concern in judging L2 speech as part of proficiency. It may well be that an accumulation of disfluency features over an extended period of time is a valid benchmark for assessing fluency: though perhaps we need a great deal longer than the ten minutes or so of most oral tests that provide no more than a snapshot of performance. The focus could be on, for example, word searches, or struggles with grammatical form, which might provide a more valid assessment of L2 fluency performance. If we were to judge the talk of the L1 teachers in the lessons from which our data come, using similar snap-shots of some sections of the lessons, the talk of these teachers could be assessed as very disfluent. What we do with learners in test situations is to take such snap-shots when they are in stressful testing situations, ask them to answer questions starting with ‘why’ or ‘how’ or produce lengthy descriptions with a heavy cognitive load. It would be unsurprising if these demands did not lead to an increase in pausing, hesitating and self-repair. Alternatively, it may be better to rely on global intuitive judgements of fluency rather than by counting instances of disfluency markers when at our current state of knowledge it appears impossible to paint a full picture of contextual variables and the conversational practices that influence the talk of speakers.

Agreement on what constitutes perceptions of fluency, and reliability of judgements, is important because of its “frequent use as a descriptor of oral performance in the course of assessment” (Chambers, 1997, p. 535), often in circumstances of great significance in the lives of learners (Derwing et al., 2004). At its most basic conception in a communicative language teaching and assessment framework, judgements of fluency are made at the discourse or utterance level, and as Eskey (1983, p. 316) observes, “utterances either do the job or they don’t.” While Eskey’s concern was with acceptance of inaccurate language

because it 'did the job', promoting fluency or addressing disfluency in the classroom entails not only clarity and agreement on the features of utterances that do, or don't do, the job, but practical measures as well. On the basis of findings such as these, Chambers (1997) concludes that qualitative analysis of temporal variables provides some useful directions for addressing fluency in second language classrooms. It may be that not pause length, but pause frequency sets learners apart from L1 speakers of a target language, that is, utterances consist of shorter word groups. The findings of researchers such as Towell et al. (1996) suggest that the link between fluency and syntactic complexity means fluency is related to proficiency in terms of syntactic and lexical knowledge that facilitates processing in real time (Chambers, 1997) to produce longer runs of speech with fewer hesitations. Indeed, many silences and hesitations and repairs can indicate high interactional competence. The problem for assessors of second language fluency is to establish which disfluency features are reflections of low proficiency, for example grammar or vocabulary, and which are reflective of proficient interactional practices. This applies as much to informal or casual assessments of fluency at job interviews for example, as to formal test situations. This may be a tall order for practitioners, given our current state of knowledge. However, for many test candidates the stakes are high. Can we afford to risk making judgements that may affect these learners' futures? We have no easy answers, but urge a re-appraisal of criteria used in fluency assessment.

References

- Bortfeld, H., Leon, S., Bloom, J., Schober, M., & Brennan, S. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech, 44*(2), 123-147.
- Brouwer, C. E. (2003). Word searches in NNS–NS interaction: Opportunities for language learning? *The Modern Language Journal, 87*(4), 534-545.
- Chambers, F. (1997). What do we mean by fluency. *System, 25*(4), 535-544.
- Clark, H. H. (2002). Speaking in time. *Speech Communication, 36*(1-2), 5-13. doi: doi:10.1016/S0167-6393(01)00022-X
- Derwing, T., Munro, M., & Thomson, R. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics, 29*(3), 359-380.
- Derwing, T., Rossiter, M., Munro, M., & Thomson, R. (2004). Second language fluency: Judgments on different tasks. *Language Learning, 54*(4), 655-679.
- Eskey, D. (1983). Meanwhile, back in the real world...: Accuracy and fluency in second language teaching. *TESOL quarterly, 17*(2), 315-323.
- Fillmore, C. J. (1979). On fluency. In C. J. Fillmore, D. Kempler & W. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 85-101). New York: Academic Press.
- Freed, B. F. (1995). What makes us think that students who study abroad become fluent? In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-148). John Benjamin Amsterdam.
- Gardner, R. (2001). *When listeners talk*. Amsterdam: Benjamins.
- Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *Canadian Modern Language Review/La Revue canadienne des langues vivantes, 61*(3), 325-353.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York: Academic Press.
- Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
- Guillot, M.-N. (1999). *Fluency and its teaching*. Clevedon, England: Multilingual Matters.
- Jefferson, G. (1984). Notes on a systematic deployment of the acknowledgement tokens “Yeah”; and “Mm Hm”. *Research on Language & Social Interaction, 17*(2), 197-216.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*(2), 145-164. doi: doi:10.1016/j.system.2004.01.001
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*(3), 387-417.
- Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25-42). Michigan: The University of Michigan Press.
- Levelt, W. J. (1989). *Speaking from intention to articulation*. Cambridge, MASS: M.I.T Press.
- MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in an L2: A situational model of L2 confidence and affiliation. *Modern Language Journal, 82*(4), 545-562.
- McCarthy, M. (2010). Spoken fluency revisited. *English Profile Journal, 1*(1), 1-15.
- Mushin, I., & Gardner, R. (2008). Silence is talk: Conversational silence in Australian Aboriginal talk-in-interaction. *Journal of Pragmatics, 41*(10), 2033-2052.
- Nation, P. (1989). Improving speaking fluency. *System, 17*(3), 377-384.

- Oviatt, S. (1994). *Interface techniques for minimizing disfluent input to spoken language systems*.
- Ranta, L., & Derwing, T. M. (2000). *Accuracy, fluency and learner style*. Paper presented at the meeting of the American Association of Applied Linguistics, Vancouver, British Columbia, Canada.
- Raupach, M. (1980). Temporal variables in first and second language speech production. In H. W. Dechert & M. Raupach (Eds.), *Temporal Variables in Speech* (pp. 263-270). Mouton: The Hague.
- Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse processes*, 14(4), 423-441. doi: DOI: 10.1080/01638539109544795
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 696-735.
- Schegloff, E. A. (1979). The relevance of repair to syntax-for-conversation. In T. Givón (Ed.), *Syntax and semantics* (Vol. 12, pp. 261-286). New York: Academic Press.
- Schegloff, E. A. (1984). On some questions and ambiguities in conversation. In J. A. J. Heritage (Ed.), *Structures of Social Action* (pp. 28-52). Cambridge: Cambridge University Press.
- Schegloff, E. A. (1990). On the organization of sequences as a source of 'coherence' in talk-in-interaction. In B. Dorval (Ed.), *Conversational organization and its development* (pp. 51-77). Norwood Ablex.
- Schegloff, E. A. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff & S. Thompson (Eds.), *Interaction and Grammar* (pp. 52-133). Cambridge: Cambridge University Press.
- Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis I*. Cambridge: Cambridge University Press.
- Schegloff, E. A. (2010). Some Other" Uh (m)" s. *Discourse Processes: A Multidisciplinary Journal*, 47(2), 45.
- 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.
- Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14(4), 357-385.
- Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, 17(1), 84-119.
- Wennerstrom, A. (2000). The role of intonation in second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 102-127). Ann Arbor: University of Michigan Press.