Measures of Fluency in Speaking
Do Hesitation Markers Show Fluency or Disfluency?

Lancaster University Graduate School
NAKAMURA Eiko

This paper explores phenomena of learners' oral language seen in the immediate task repetition of Poster Carousel. Close examination of features usually considered as disfluency markers to measure fluency reveals different functions in different situations. The phenomena such as unfilled/filled pauses and other lexical fillers (e.g., repetitions, self-repairs, prefabricated chunks, and reformulations) in transcribed speech data are qualitatively analyzed and classified into two categories: those related to the speaker's concerns for own performance and those concerning the interlocutor's understanding. The functions of these phenomena are seen to change over successive cycles of the task. Learners' use of these devices for own performance is seen to decrease while that for the interlocutor's understanding increases. Most importantly, the shift of phenomena that emerged in oral interaction through task repetition reflects how learners overcome their language problems. The results of measuring phenomena in different ways are also compared in the data.

1. Introduction

"Are errors really a problem, or are they an important part of learning itself?"
(Allwright and Bailey, 1991: 83)

In communicative approaches using tasks in language pedagogy, phenomena in learners' speech have been used as tools to measure their language competence or are often considered as problems to solve for learners to develop their target language. However, do we know for certain that those specific phenomena are really problematic behaviors? Or is it possible that the very phenomena could be used as tools for learners to develop their target language? Understanding the phenomena about how learners produce them in speaking their target language can be a key to find a better pedagogical approach. This paper tries to seek for the best way to see learners' language, focusing on fluency measures and how spoken fluency should be analyzed.

2. Literature Review

In research on language learning and use, pauses are often used as measures of fluency by the length of pauses (Raupach, 1987; Mehnert, 1998; Kormos and Dénes, 2004) and by the number of pauses per c-unit, per t-unit, or per minute (Mehnert, 1998; Skehan, 2001; Bygate, 2001; Tavakoli and Skehan, 2005). Another way to look at pauses is the location of pauses. Butterworth (1980) believes that micro-planning is made at clause boundaries for lexical selection, which leads to

Hesitation markers such as repetitions and self-repairs are also considered when measuring of fluency. Skehan (2001) takes reformulation, replacement, false starts, and repetition as disfluency indicators. However, Shehadeh (1999) sees these markers caused by output difficulties as self-initiated clarification attempts, instead of disfluency. Some research focuses on self-corrections or repairs in learners' language for various purposes, not just to measure disfluency, to see the difference between proficiency levels (Kormos, 2000), problem-solving mechanism (Dörnyei and Kormos, 1998), and self-repairs through task repetition (Lynch and Maclean, 2001).

Hesitations also indirectly influence other measures of fluency. Some research includes hesitation markers (filled pauses, repetitions, and self-repairs) in the number of words/syllables to measure fluency (Dörnyei, 1995; Towell et al., 1996; Kormos and Denes, 2004), while other research excludes them (Mehnert, 1998; Ortega, 1999; Foster et al., 2000; Skehan, 2001; Bygate, 2001; Tavakoli and Skehan, 2005).

The various ways of looking at language use seem to have caused this phenomenon, as Foster et al. (2000) note that "different researchers may wish to deal with actual linguistic material within the false starts and corrections in different ways, depending on their interests" (p. 368). When we see language ability as language knowledge, these hesitations in learners' language use are typically seen as time-creating devices to search for words/form not yet automated. Therefore, these hesitations, such as repeated words or part of a word, incomplete words before repaired, and filled pauses, are not computed as part of articulated words.

On the other hand, the use of time-creating devices can be considered to be a type of communication strategy. Dörnyei (1995) claims:

... we were particularly interested in one aspect, the ability to fill the time with talk, which contrasts with a characteristic feature of L2 speech ... in which the learner keeps grinding to halt, pauses for lengthy periods, and often gets so lost that the interlocutor loses patience, or a complete communication breakdown occurs. In measuring speech rate, fillers, lexicalized hesitations ... and repetitions are considered to be part of fluent speech ... (p. 71)

Dörnyei examined how teaching communication strategies affects the speech rate, which includes these features as fluency markers. A study of comparing results between including and excluding these markers in the number of words/syllables to examine fluency is needed. Whether these hesitation markers are computed as part of words/syllables or not, they are likely to be uniformly computed as either positive or negative features of fluency. Hesitations, however, could have various functions in communication.

Another way of seeing fluency might be by how many accurate and target-like collocations an L2 speaker can produce through on-line planning. Using chunks reduces an L2 speaker's cognitive burden from paying attention to both the form and meaning of language. Forster (2001) reports that non-native speakers are processing language more through rules than routines, when compared with native speakers. The use of prefabricated chunks (one type of collocations) as a time-creating device instead of hesitations would be a sign of fluency (Dörnyei, 1995).
On the contrary, Fulcher (2003) has doubts about counting constituents to measure fluency. He argues "the initial problem that emerged from 'counting' pauses or repetitions stemmed from the fact that the number of pauses did not automatically translate into a perception of reduced fluency" (p. 99). He examined speech data through discourse analysis and found several types of pauses in different situations: "end-of-turn pauses," "content-planning hesitation," "grammatical-planning hesitation," and "addition of examples, counter-examples, or reasons to support a point of view" (pp. 100-101). Fulcher related the use of these types of pauses with examinees' proficiency levels. Low proficiency examinees are not always producing more pauses than higher proficiency examinees: both examinees in level 1 and level 5 used end-of-turn pauses, though the reasons of the use are different: content planning hesitation increases as the level goes up to level 4, but in level 5 it goes down to even under level 1; and Level 2 and 3 have both grammatical-planning hesitation and addition of example, counter-examples, or reasons to support a point of view; but the latter is not likely to occur in low proficiency learners' language. Just computing pauses or pausing time is not likely to be a perfect way to measure fluency.

Though Fulcher's sample was small, these four types of pauses are also seen in my previous study (Nakamura, 2008a). Just computing pauses/pausing time or hesitations is not likely to be a perfect way to measure fluency. A new way to distinguish pauses/hesitations as disfluency markers from others is needed. Understanding the complex emergence of phenomena in learners' language output could be a key to a deeper understanding of language learning.

3. Research Questions
1. Do unfilled/filled pauses and hesitations such as repetitions and self-repairs seen in learners' language show fluency or disfluency?

2. Is fluency evaluation different between measures including or excluding filled pauses and hesitations?

4. Study
In my previous study (Nakamura, 2008a) pauses are seen to have different functions in different situations, not just as markers of disfluency. This study deepens the previous study by closely examining learners' speech data, focusing on measures of fluency.

4.1 Participants
Participants are six pairs of Korean and Japanese students. They are six male Korean students and six female Japanese students studying at a Japanese university. Their language proficiency levels are between pre-intermediate and pre-advanced (e.g., STEP test: pre-2nd, 2nd grade, TOEFL: 170-210, TOEIC: 500-600). In addition, four Japanese participants have experience of study abroad (two-four weeks) and one Korean student has experience of living in English speaking country for four years.

4.2 Task
The task used in this study is a modified Poster Carousel task. The theme of the task is "Korean
and Japanese Cultures”. Before starting the task, six pairs of Korean and Japanese students are formed and each participant makes a poster, based on the person’s culture, on the same topic of their partner’s in advance. The topics are sports, games, movies, food, and so on. Initially, each partner by turn explains his/her poster to the other person in the pair (Cycle 1). Then, Japanese students visit other posters asking questions one by one, while Korean students stay at their base to answer visitors' questions (Cycle 2-6). After Japanese students come back to their base, they reverse the roles of host and visitor. For the second round Korean students visit the posters, while Japanese students stay at their base as host (Cycle 2-6). Each cycle (visit) is five minutes for the interaction between visitor and host, except for the first cycle (10 minutes), in which partners explain their poster to each other (Cycle 1). After finishing the task, they fill out a questionnaire and discuss the merit of doing this task.

4.3 Procedure
First, the nature and the functions of unfilled/filled pauses and other hesitation markers in transcribed speech data are qualitatively examined in learners’ turns of over 10 words or over 20 seconds on the same topic over successive cycles of the task. Second, hesitation markers are classified into two categories: those related to the speaker’s concerns for own performance, such as filled pauses, repetitions, and self-repairs, and those concerning the interlocutor’s understanding, such as repetitions as understanding response, reformulation to help the interlocutor understand, non-lexical feedback to the speaker (e.g., uh-huh). Lastly, speech rates (the number of words per second) and pause/time ratios (percentage of overall pausing time) are quantitatively examined to see the overall view. Speech rates including/excluding hesitation markers are also compared with line graphs and tests.

5. Analysis
The following are analyses of pauses and hesitations in transcribed speech data. Two students’ language speech samples over successive cycles of the task are closely examined and analyzed.

Notes: (number): pausing time, **encircled**: filled pauses, **bold** letters: repetitions, **bold italics** self-repairs, **shadow**: prefabricated chunks, **shadowed** bold: understanding response, **encircled** shadow: non-lexical feedback, **shadowed bold italics**: reformulation for the interlocutor, >speak quickly<, SR: speech rate, P/T: pause/time ratio, OP: hesitations for own performance/total time, IU: hesitations for the interlocutor’s understanding/total time.

5.1 Example 1: Sun
The first example shows a Korean student Sun’s speech samples over the cycles. Sun is talking about FIFA Online, a computer game in Korea. Sun produces lots of unfilled/filled pauses, and repetitions in the first interaction. The explanation is not clear and hard to understand.

*Excerpt 1: Interaction with the first visitor in Cycle 1*

1 Sun: **Uh**. (1) **uh** Korean mode is (1) in ma (2) **uh** Korean mode is (2) it is na
2 **Korean mode** is match mode **no** (2) hantai.
3 Sae: **Hantai? (opposite?)** Hhhh
4 Sun: "Uh Korean mode means it means I manage my team.
5 Sae: Shie?
6 Sun: my my team and (1) I take Koshien and I manage my team then then means Korea. (2) Korean means work.
8 Sae: Work?
9 Sun: Work. Work is my work means managing team and
10 Sae: Uh huh
11 Sun: match match mode is (1) also playing without difference.

[43 words, 52 sec. SR: 0.83, P/T: 12/52=0.23, OP: 16/52=0.40, IU: 2/52=0.038]

With the second visitor his explanation becomes much clearer and easier to understand. The pause/time ratio drastically decreases from 0.23 to 0.08 and the speech rate increases from 0.83 to 1.45. Hesitations for own performance also drastically decrease while those for the interlocutor's understanding increase (lines 17, 21).

Excerpt 2: Interaction with the second visitor in Cycle 2

12 Nao: what can I do? What must can must =
13 Sun: = (1) yeah, first, you link link to this page, link this homepage =
14 Nao: = Homepage, uh.
15 Sun: >And you have to take the< (1) use sign and (1) menu log in log in =
16 Nao: = Log in, uh.
17 Sun: log in this game, and download this game, and log log in, and play this game.
18 Nao: Huh huh huh. (1) it is need to to money? [Need money?]
19 Sun: Ah, it needs no money.
20 Nao: Oh, free?
21 Sun: It's free game.

[135 words, 93 sec, SR: 1.45, P/T: 7.5/93=0.08, OP: 7/93=0.08, IU: 11/93=0.12]

In Excerpt 3 in Cycle 4, no unfilled pauses are seen and he is explaining smoothly here. However, one problem occurs in this cycle. His interlocutor does not understand his pronunciation flay and asks him to repeat (line 27). He hesitates and produces a filled pause (line 28).

Excerpt 3: Interaction with the fourth visitor in Cycle 4

22 Sun: Eh, this is a soccer games.
23 Non: Soccer game. [FIFA.
24 Sun: [Playing online.
25 Non: Ehu? Eh, wha, eh, what what what do you do th on this game?
26 Sun: Yeah, I connect to this game and flaying with computers or flaying other flayers.
27 Non: (1) moikkai (one more time).
28 Sun: Eh. If I connect this game, I can play with computers or ma another flayers.
29 Non: Um. Fle flayer?

[27 words, 18 sec, SR: 1.50, P/T: 0, OP: 2/18=0.11, IU: 10/18=0.56]
In Cycle 5 he seems to have remembered the problem caused by his pronunciation of 'for p. His hesitations (repetitions and a filled pause in lines 30) occur before using the problematic word. Surely again his interlocutor does not understand his pronunciation of 'flay and asks him to repeat (line 32). He tries to explain without using 'flay after the trouble. His interlocutor finally understands the problematic word 'flay by explaining it with the word choose (lines 36, 37).

Excerpt 4: Interaction with fifth visitor in Cycle 5

30 Sun: If you connect this game and if you connect this game, eh you 'flay with other
31 players.
32 Kei: Eh? Mo Sorry?
33 Sun: Ah you download this game, [install this game, and you login login to this game,
34 Kei: [Ha, this game, yeah,]
35 Sun: you can 'flay with other people.
36 Kei: [Meh, yo:u are choosing a team or what? You, hhha, nanteiebaï (what can I say?)
37 how you can enjoy this game? You () can choose eh team members from =
38 Sun: = Yeah, I choose my team.
[39 words, 23 sec, SR: 1.70, P/T: 0, OP: 6/23=0.26, IU: 1/23=0.04]

In Cycle 6 he uses the word choose to explain the same thing before using the problematic word 'flay (line 42). He seems to have found a solution to the problem. He produces one unfilled pause and a filled pause after saying “flayed this game”, which is located before changing the topic (line 46). He may be waiting for his interlocutor to understand his pronunciation of 'flayer or thinking of a new example “Winning Eleven”. Either way, these hesitations are obviously not caused just by lack of language knowledge.

Excerpt 5: Interaction with sixth visitor in Cycle 6

39 Yu: So you (1) nanteiyuundaro (what can I say?) you control the player or
40 Sun: team.
41 Yu: to ah team. You may o th your
42 Sun: I ch I can choose one team, real real team.
43 Yu: Oh, really?
44 Sun: Yeah.
45 Yu: [Eheh.
46 Sun: And I flayed this game. eh (1) In PlayStation 2 Winning Eleven, do you know Winning
47 Eleven?
48 Yu: [Yeah, I know.
49 Sun: !This is almost same.
[29 words, 17 sec, SR: 1.71, P/T: 0, OP: 3/17=0.18, IU: 3/17=0.18]

5. 2 Example 2: Seo
The following example shows another Korean student Seo's speech samples. Seo is in the advanced level with the experience of living in the US for four years when he was in elementary
school. He tends to use prefabricated chunks (lines 50, 52, 55, 57, 59) instead of other devices. He also frequently reformulates his expressions, which are often accompanied by his interlocutor's non-lexical response (lines 50-52, 53, 59-61, 63-65). He is possibly reading his interlocutor's facial expressions when he reformulates. He also repairs forms and recycles reformulations in the different cycles (lines 50→55, 53→75-76).

Excerpt 6: Interaction with the second visitor in Cycle 2

50 Seo: Well, the biggest difference is that (1) [they don't have leg: they don't have legs: Hah!]
51 Kei: [laughs]
52 Seo: attacks but this taekwondo is, like kind of, they have eh they attack with both hands and legs, but usually legs [they use legs to attack: Hec-]
54 Kei: [laughs]
[39 words, 26 sec. SR: 1.50, P/T: 1/26=0.04, OP: 6/26=0.23, IU: 15/26=0.58]

Excerpt 7: Interaction with the third visitor in Cycle 3

55 Seo: [mm, well, first of all, judo doesn't use legs =
56 Yu: = Oh, really =
57 Seo: = Yes, [ah, eh] they are like () when they tackle they use legs but ()
58 Yu: [laughs]
59 Seo: ah:hu, what can I say, taekwondo [they have] is kind of kicking skills, they have
taekwondo [laughs]
60 Yu: [laughs]
61 Seo: kicking skills
62 Yu: So they don't use hands?=
63 Seo: = [They use hands, but usually they use legs. Their legs are much more]
64 Yu: [laughs]
65 Seo: [stronger than hands] hhaa.
[40 words, 24 sec, SR: 1.67, P/T: 1/24=0.04, OP: 4/24=0.17, IU: 24/24=1]

His speech rate goes up from Cycle 2 to Cycle 3 (1.50 to 1.67), but down in Cycle 5 (1.45). One big difference seen here is the interlocutor’s reactions. The interlocutor in Cycle 5 often cuts in to ask him questions during his explanation (lines 67, 72). He seems to wait for his interlocutor to understand him, which is displayed in the distribution of his unfilled pauses. His unfilled pauses are usually located at the sentence-junctures, while those in Cycle 5 are located at phrase-junctures (lines 68, 70). Phrase-juncture pauses are thought to be for lexis searching (Butterworth, 1980). However, his phrase-juncture pauses are obviously not to search for words/form or content because those expressions have already been repeatedly used in the previous cycles.

Excerpt 8: Interaction with the fifth visitor in Cycle 5

66 Seo: Oh, really, [ah:hu] () this is a kind of traditional sport =
67 Nao: = in Korea? =
68 Seo: in Korea. Ah: huh, well (1) You wear similar uniforms (1) like judo
69 Nao: Ha-ha.
70 Seo: but ah: huh, well, () the biggest difference () from judo is (1) () use foot, they can
71 use foot =
72 Nao: = foot =
73 Seo: = to kick the other person. Hhhhh.
74 Nao: Hhhhh. It's aggressive. Hhhhh.
75 Seo: Hhhhh. It is. Eh and so () they can use both hands and both legs, but () usually ()
76 they use legs to attack.
[96 words, 66 sec, SR:1.45, P/T: 9/66=0.14, OP: 18/66=0.27, IU: 25/66=0.38]

6. Results and discussion
The features classified into two categories of hesitations (for own performance and for the
interlocutor's understanding) are also quantitatively examined. Table 1 shows speech rates of six
host students' language output of over 10 words or 20 seconds per turn on the same topics. The
t tests of the speech rates between the 1st and other cycles show significant difference (1st-2nd:
p<.05, 1st-5th: p<.01). However, the pause/time ratio does not show any statistical difference, which
is the opposite result to my previous study (Nakamura, 2008b). In the previous study the
pause/time ratio showed significant difference but speech rate did not. This could be caused by the
size of data, since speech data were examined only on the same topics in this study. On the other
hand the same topics, which possibly gave learners more opportunities to repeat the similar
expressions, could have affected learners' speech rates.

Table 1. Paired t tests (two-tail) Speech rates over successive cycles

<table>
<thead>
<tr>
<th>Pairs</th>
<th>Mean</th>
<th>S.D.</th>
<th>t Value</th>
<th>df</th>
<th>sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle 1</td>
<td>.79</td>
<td>.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cycle 2</td>
<td>1.20</td>
<td>.36</td>
<td>-2.902</td>
<td>5</td>
<td>.034</td>
</tr>
<tr>
<td>Cycle 1</td>
<td>.82</td>
<td>.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cycle 5</td>
<td>1.43</td>
<td>.38</td>
<td>-4.786</td>
<td>4</td>
<td>.009</td>
</tr>
</tbody>
</table>

Figure 1 shows learners' speech behaviors over the cycles. The use of hesitations for own
performance drastically decreased in the second cycle while those for listener's understanding and
prefabricated chunks do not show a clear movement in the quantitative examination.

Figure 2 is the comparison of speech rates including/excluding hesitations. Different from my
prediction, the three lines show a similar curve and all the t tests between 1st and other cycles in
three ways of measuring speech rates show significant difference (p<.05). This result supports the
research by Kormos and Denes (2004), who examined the relation between these features and
teachers' perception of fluency. In their research the number of unfilled/filled pauses and other
disfluency phenomena were not found to influence teachers' perceptions of fluency.

Hesitations are also seen to have both functions of fluency and compensation for dis fluency in
pre-intermediate to pre-advanced proficiency learners' language. One way of distinguishing
pauses/hesitations of disfluency from fluency might be to distinguish those related to the speaker's concerns for own performance from those concerning the interlocutor's understanding.

7. Conclusion
The result indicates that fluency evaluation is not different between including/excluding hesitation markers as part of words in pre-intermediate and pre-advanced learners' language.

This study has focused on the functions of pauses and hesitations in individual learners' language in a Poster Carousel task and shows pauses and hesitations are natural phenomena in speaking. It also has tried to find relevant measures in order to establish a framework for distinguishing features of fluency from disfluency: those related to concerning own performance or the interlocutor's understanding.

There might be some keys to improve learners' language in speaking by shifting the speaker's attention from concerning for own performance to concerning for interlocutor's understanding.

Acknowledgments
I would like to thank Martin Bygate for guidance, Ian Nakamura for useful comments on an earlier draft of this paper, and reviewers for helpful advice for this revision.

References


learning, teaching and testing (pp. 23-48). Harlow: Pearson Education.