This study sought to develop a description of second language fluency, based on the concept that the microcontext of the testing task and the resulting psycholinguistic demands on speech production affect the nature of the discourse produced so that different tasks have differential impact on an individual's display of fluency. It investigated how assessment of learners' fluency varies with four speech task structures (cued dialogue, uncued dialogue, cued monologue, uncued monologue) and any difference in oral discourse produced in performance of the four tasks. Subjects were 46 students in a private school for English as a Second Language in Brazil, divided into three proficiency levels. Each performed one oral task in Portuguese and five in English (two dialogues, two monologues, and a cued monologue from a standardized test), filled out a questionnaire, and was interviewed. Four raters assessed performances on scales for continuity/flow (length of utterance), ease of expression (quality of pause or hesitation), and rate of speech (speed and evenness). As hypothesized, task structure and interactivity both influenced speakers' fluency ratings and discourse variability, with greater fluency apparent in dialogues. Implications for theory of fluency, instructional and testing practice, and research are considered. Contains 61 references. (MSE)
The Role of Task Structure in Oral Fluency Assessment

Roseli Ejzenberg
Montgomery College

1. Introduction: Proficiency versus Fluency

Although interest in language testing dates back to World War I and decades of researchers have addressed the question of defining and measuring nonnative oral proficiency, (i.e., overall language ability) and fluency, as Freed (1995) points out, these constructs are “often interchanged ... and subject to multiple and overlapping definitions.” This paper addresses several pressing problems in the field: How should we test speakers at varying levels of oral proficiency in second/foreign language? What are we really testing at these levels—proficiency, fluency, or strategic competence? How are these three related? And the most fundamental question, What exactly is nonnative oral fluency? More specifically, the study investigates whether the display of nonnative oral fluency is context-dependent and, if so, whether there are quantitative and qualitative differences in the oral discourse produced by very fluent and very disfluent speakers.

Fluency is here defined as a component of language proficiency that is present with different degrees of saliency across proficiency levels becoming more salient as a speaker moves towards higher proficiency levels. Whereas a novice or a less proficient speaker is still “groping for words” (Foreign Service Institute fluency scale cited in Vallette 1977:159) and struggling to get a message across, a more advanced speaker will be concerned about how fluently that message will be conveyed. This poses the question of whether fluency is a “mysterious quality that appears at a certain level of oral proficiency” (Riggenbach 1989:156).

Despite the crucial role that fluency plays as a major component of foreign language proficiency (Graham 1978; Mullen 1978; Oller 1989), its relationship with it is still unclear (Butler-Wall 1986; Hieke 1985). Little research has applied findings from other research disciplines to the second/foreign language context, few empirical studies have examined the nonnative fluency construct (Butler-Wall 1986), and there is little understanding of what exactly comprises fluency (Riggenbach 1989).

Native and nonnative speaker oral fluency—variously defined as “speech which is effortless and smooth” (SPEAK Test rating scale in Educational Testing Service 1987), or the “ability to plan and arrange ... speech performance [and] to avoid numerous very long pauses” (Raupach 1980:268)—has been the object of study in several subfields of research. Clearly, these definitions vary considerably in the scope of the phenomenon because fluency has been described at different levels of occurrence, namely: at the utterance level (Hieke 1988) versus at the discourse level (Goffman 1981; Jefferson 1979); at the phonetic (Hieke 1984a; Jull 1987) and grammatical levels (Murphy 1986) versus the semantic (i.e., information flow such as in...
Hieke 1984b and Schaefer 1980) and the sociolinguistic or functional-interactional levels (Kramsch 1986).

These overlapping definitions resulting at times from a lack of crossreferencing among the several fields of research (Butler-Wall 1986) has yielded a confusing state of affairs, which is symptomatic of a shortcoming in the way fluency-related phenomena have been addressed in many studies. First of all, a full account of fluency requires an analysis of its occurrence at both the utterance and the discourse level because fluent performance at the utterance level may be disfluent at the discourse level (See, for example, Goffman 1981, Jefferson 1979, or Butler-Wall 1986:284 for a discussion of the adverse effects of “relentless” or hyperfluency).

In addition, a sociocommunicative, sociocognitive view of language (Langer 1987) posits that discussions about language need to address the context of language production (Bachman 1990:111–112; Cathcart 1986; Finocchiaro and Brumfit 1983:13; Hymes 1972; Labov 1972). Even though second language researchers are increasingly sensitive to the effect of task or elicitation procedure on speech production (Crookes 1990:49), as has been reported in the extensive literature on the topic (see, for example: Doughty and Pica 1986; Duff 1986; Long 1980, 1981, and 1983; Porter 1986; Schachter 1986; Young 1984), it seems that fluency research has lagged somewhat behind in this respect.

Finally, with the recent questioning (Greene and Cappella 1986; Hieke 1984a, 1984b, 1985; Postma, Kolk, and Povel 1990) of early fluency-related research (Chafe 1979; Goldman-Eisler 1968; Henderson, Goldman-Eisler, and Skarbek 1966), there is an urgent need to re-conceptualize fluency and the methods used to investigate and test it. These issues are especially pressing in the testing field where the quest for more authentic and more satisfactory tasks for the direct and semi-direct assessment (Clark 1986) of speaking has assumed focal interest (Bachman 1990; Weir 1990).

2. Synthesis of Related Research

The present study sought to offer a better description of what fluency is being grounded on the hypothesis that both its framework of occurrence, here understood as the microcontext of a testing task, and the resulting psycholinguistic demands on speech production affect the nature of the discourse produced so that different tasks have a differential impact on a participant’s display of oral fluency.

A very selective synthesis of eight main research findings about the fluency construct as they pertain to the specific variables under study follows (For comprehensive literature reviews, see Ejzenberg 1992 and Koponen 1992. For a good synthesis, see Freed 1995).

1. Nonnative speech seems to reproduce native speech organization but with increased pause and hesitation phenomena, decreased speech rate or rate of articulation, and a “loss of fluency” (Deschamps 1980; Raupach 1980 cited in Olynyk, D’Anglejan, and Sankoff 1990).
2. Fluency is a potentially strong indicator of oral proficiency, specially when testing more proficient speakers (This an outgrowth of Riggenbach's research). The fluency subscale was "the clearest predictor of overall proficiency as measured on this FSI interview, and thus could be the most reliable of the subscales."(Hinofotis, Bailey, and Stern 1981 quoted in Riggenbach 1989:157).

3. Fluency has been shown to be sensitive to context (Ochs 1979 in Olynyk, Sankoff, and D'Anglejan 1983:216) and to be differentially displayed accordingly (Riggenbach 1989). Olynyk et al. (1983), for instance, examined fluency in planned versus unplanned discourse and Brown and Yule (1983), who studied communicative stress, found that a speaker will do best-- that is, perform most fluently-- under conditions of least communicative stress. Three conditions contribute to higher or lower degrees of communicative stress: features of the context, state of knowledge of the listener, and type of task, including status of knowledge and task structure. Task structure, operationalized as cuing (cued versus uncued condition) and task interactivity (dialogue versus monologue condition), was one of the variables under consideration in the study.

4. Cuing, also called "external support" or "structure" to a task (Brown and Yule 1983:36), is an important variable in determining task structure so that finding the appropriate balance in the amount of cuing inserted in the input will determine the speech outcome of the testing situation. It is helpful to the speaker if the information in the task provides its own structure so that the language is externally supported by the requirements of the task. Thus it is easier for any speaker to give an account of a series of events than it is to provide an argument for 'why' those events occurred in that order. However, too much cuing may stifle production and prevent speakers from freely structuring their discourse (Coulthard 1985:3-4). And in testing more proficient speakers, it is desirable to deliberately insert tasks with increased communicative stress by, for instance, asking a speaker "to create a structured account of an event which has no obvious structure." (p. 36) In contrast, tasks that fall short of supplying enough cuing can also bias the testing performance so that the examinee may be mistakenly perceived to be less fluent than s/he really is.

5. The amount of communicative stress present in a task that demands a display of fluency in short turns or only one or two utterances such as the Oral Proficiency Interview is much less than in a task that demands display of fluency in long turns, i.e., "a string of utterances" (Brown and Yule 1983:16). This distinction refers expected speech response at the utterance level versus at the discourse level. Clearly, "short turns do not demand much of the speaker in the way of producing structure." (p. 16). On the other hand, "as soon as a speaker 'takes the floor' for a long turn, tells an anecdote, tells a joke, explains how something works, justifies a position, describes an individual, and so on, he takes responsibility for creating a structured sequence of utterances which must help the listener to create a 'coherent' mental representation of what he is trying to say" (p. 17).
6. Highly fluent speakers share all or most features of fluency while nonfluent speakers will be disfluent in idiosyncratic ways (Olynyk et al. 1990:139–140; Riggenbach 1989:164).

7. An increase in frequency of occurrence of hesitation phenomena correlates with the creation of novel utterances and with situations of greater uncertainty such as cognitively more demanding tasks (Goldman-Eisler 1968; Ochs 1979). Conversely, speech rate increases when a speaker is producing rehearsed or repetitious speech (Ejzenberg, in press) or is engaged in situations of greater predictability such as cued tasks.

8. Languages differ in the patterning of hesitation phenomena and fluent speech (Butler-Wall 1986). Although fluency and pause profiles are idiosyncratic, they do vary systematically across linguistic communities (Scollon and Scollon 1981, 1983) and even within the same individual across languages (Olynyk et al. 1990:141). The present study was conducted with a homogeneous linguistic group whose native language is Portuguese, and systematic variability in their nonnative English speech production was expected. For example, compared to Americans, educated Brazilian adults are less encumbered by strict rules governing the appropriateness of interruptions and overlap in informal conversation.

3. Research Questions and Hypotheses

Three variables were investigated in this study:

- Task structure, operationalized as cuing (cued versus uncued condition) and task interactivity (dialogue versus monologue condition);
- Display of fluency, operationalized as fluency ratings obtained in the performance of four different speaking tasks (cued dialogue, uncued dialogue, cued monologue and uncued monologue); and
- Fluency markers: speech features that differentiate the speech of the very fluent and the very disfluent.

The research questions addressed were:

1. Does assessment of learners' fluency vary with respect to four dimensions of task structure, namely, cued dialogue, uncued dialogue, cued monologue, and uncued monologue? If so, how?

2. Is there a difference in the oral discourse produced in the performance of the four tasks? If so, what are the discourse features that mark the difference between fluent and disfluent speech? Are these discourse markers influenced by the context of speech production (i.e., task and task structure)? If so, how?
It was hypothesized that task structure, operationalized as task interactivity and cuing, would affect speakers’ discourse variability and fluency ratings. Both the dialogue and cued conditions were expected to provide less communicatively stressful contexts for oral production so that speakers’ fluency ratings on those tasks would be higher than on the uncued and monologue conditions. In addition, it was presumed that task structuring would have an impact on discourse variability of fluency-related features and that there would be a relationship between such variability and the fluency ratings obtained on the tasks.

First, compared to the dialogues, monologues were expected to lend themselves to more nonnativelike pause and hesitation phenomena and to a clearer picture of the interlanguage level of the speakers who would be required to display fluency at the discourse level rather than only at the utterance level. A number of possible scenarios were projected: Examinees’ speech rate, in monologic discourse, would presumably be slower. Since faster speech rate is an important marker of fluent speech, speakers’ performance in the monologue tasks would be rated as less fluent.

Second, cued tasks were expected to ease speech production by alleviating the speaker’s burden of structuring the discourse and, consequently, trigger the production of more fluent speech so that speakers’ performance in them would be perceived and rated as more fluent than on the uncued tasks. It was hypothesized that the uncued dialogue would generate more negotiation of meaning and competition for turns than its cued counterpart—where both parties were aware of the directives to be accomplished in the interaction—with those more in control of the conversation being rated higher in fluency. However, it was foreseen that, for more fluent speakers, cuing might end up stifling production, especially for those at higher levels of fluency who might have competing ways of organizing their discourse.

4. Description of the Study

Research Site and Participants

The study was carried out in the fall of 1991. The site was a private English as a Foreign Language school located in Sao Paulo, Brazil. Fifty students were randomly picked from those who volunteered, and the sample was stratified (Brown 1988:112) by course level bracket which served as a reliable proxy for proficiency level. Since four cases had to be deleted from the analyses, the working sample included 46 valid cases grouped into three proficiency level brackets, namely: low (n = 10), mid (n = 18) and high (n = 18). Participants were 62% female and 38% male with a mean age of 28 and a modal age of 20, thus, a predominantly young adult sample.

Data Collection

Fifty recording sessions using a tape recorder were conducted on school premises with only participant and researcher present. Sessions consisted of one task performed in Portuguese and
five tasks performed in English (namely, two dialogues and two monologues in balanced sequence and a cued monologue extracted from the SPEAK test). Before leaving, participants filled out a short feedback form in Portuguese. By the end of the recording phase of the study, working data collected included 276 audiotaped speech samples (46 participants performing six tasks each), 49 written participant feedback questionnaires, and 49 semi-structured interviews.

**Rating the Speech Samples**

In the first rater training sessions, the four raters were introduced to the fluency scale and rubrics by using speech samples obtained in the pilot study. There were three scales, shown in Figure 1: continuity or flow indicated by length of utterances; ease of expression indicated by debilitating versus non-debilitating pause and hesitation; and rate of speech indicated by speed and overall evenness versus variability. The raters gave a holistic rating of the students' fluency on each task using the Fluency Rating Descriptors and a six-point scale ranging from 1 (very low fluency) to 6 (very high fluency). Raters were first asked to blindly rate all samples of the Portuguese task, which was a cued monologue. Then the fluency tasks were submitted to blind rating in the order: (1) cued monologues, (2) uncued monologues, (3) cued dialogues, and (4) uncued dialogues.
Interrater reliability was calculated. Examination of the alpha coefficients showed that there was very high interrater reliability (from .95 to .98). That is, comparatively speaking, the four raters scored the speaking tasks in similar and consistent ways.

*Case Studies of High and Low Fluency Speakers*

Upon completion of the rating sessions, the 46 participants were ranked according to performance on each task and then based on their overall fluency scores. In other words, the ratings
assigned by the four raters to each task performance were averaged to form each participant’s partial fluency rating on each task. Then each participant’s four partial fluency ratings (i.e., the mean of the four judges’ ratings for each task) were added and averaged into one overall fluency rating per participant. Finally, participants were rank ordered with respect to their overall fluency ratings in a distribution going from highest to lowest. After two sessions we managed to agree on the three homogeneously lowest and the three homogeneously highest fluency participants.

5. Analysis of Assessments: Fluency as a Function of Task

Both the scores and the transcribed discourse from the recording sessions were subjected to analyses to test the hypotheses of this study. First, in this section, quantitative analyses of raters’ assessments will be reported in order to investigate whether assessment of learners’ fluency varied across task. In the following section, an analysis of learners’ discourse will be reported in order to investigate the differences between fluent and disfluent speech.

Quantitative analyses of scores first addressed the issue of task validity; that is, whether the four tasks were really measuring what they purported to be measuring. The question was whether the four tasks required participants to produce a similar linguistic behavior (i.e., oral fluency) in ways that were internally valid (i.e., measuring the same phenomenon). In statistical terms this question was posed as follows: For the entire sample of 46 participants and the corresponding 276 tasks, were the fluency ratings per task significantly correlated?

As the data in Table 1 indicate, there was task validity in that all tasks were to a reasonable extent measuring what they were purported to be measuring: oral fluency. Pearson product-moment correlations among the four fluency tasks were all significant (p < .01), and ranged from $r = .62$ to $r = .79$.

Table 1: Intercorrelations Between Raters’ Assessments on Four Tasks and SPEAK

<table>
<thead>
<tr>
<th>Task</th>
<th>Cued monologue</th>
<th>Uncued monologue</th>
<th>Cued dialogue</th>
<th>Uncued dialogue</th>
<th>SPEAK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cued monologue</td>
<td>--</td>
<td>.68</td>
<td>.70</td>
<td>.72</td>
<td>.66</td>
</tr>
<tr>
<td>Uncued monologue</td>
<td>--</td>
<td>.62</td>
<td>.68</td>
<td>.69</td>
<td></td>
</tr>
<tr>
<td>Cued dialogue</td>
<td>--</td>
<td>.79</td>
<td>.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncued dialogue</td>
<td>--</td>
<td>--</td>
<td>.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPEAK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Intercorrelation mean .69 ($N = 46$)
If the comparative fluency measure extracted from the SPEAK test is included, the range of correlation coefficients is increased in the lower end (from $r = .59$ to $r = .79$). Overall, the lowest correlation between task scores was between the cued dialogue and the SPEAK task ($r = .59$) and the highest was between the cued and uncued dialogues ($r = .79$). Since the overlap was not complete—i.e., the correlation figures did not account for 100% of variability in the scores—it may be that the tasks had tapped different aspects of oral fluency or measured oral fluency to different degrees.

Once it was established that the tasks shared commonality in their purported goal of measuring oral fluency, the next step was to examine more closely the extent to which they varied in doing so. It had been hypothesized that the speaking tasks would have an impact on the speakers' display of oral fluency and that there would be variability in the fluency ratings obtained by the speakers as a result of the specific context of speech production. The tasks had been anticipated to be in order of increasing difficulty: cued dialogue, uncued dialogue, cued monologue, and uncued monologue. The uncued monologue was thus expected to provide the clearest assessment of a speaker's stage of interlanguage development because s/he would be required not only to display fluency at the discourse level rather than at the utterance level but also to organize and structure the narrative without any external support from task cues or an interlocutor's help.

Task difficulty reflects the percentage of examinees who get the right answer on a test item (Anastasi 1988:203). In the study, a correct answer was a mid to high fluency rating (i.e., 4.34 to 6.00 points). Results showed that the most difficult task was the uncued monologue (21.7%) which was followed in decreasing order of difficulty by the cued monologue and the SPEAK task (30.4%), the cued dialogue (37.0%), and the uncued dialogue (50.0%). At the other end of the scale, the easiest task was the Portuguese measure (91.7%). The calculation of the discrimination index for each task was done by using the participants who received a high global high fluency rating as the high group and using the participants who received a low global fluency rating as the low group. The results clearly indicated that, with the exception of the Portuguese measure, the tasks effectively discriminated between high and low fluency speakers. Surprisingly, the cued dialogue emerged as the task with highest discriminating power. However, given the small $N$ used in this study, these results should be viewed with caution.

Examination of the partial Cronbach's alpha coefficients resulting from deleting one task at a time indicates that no task should be deleted because the overall alpha would go down. Both the overall and the partial intertask reliability coefficients for the fluency scale without the SPEAK task are lower. In sum, the four fluency tasks in conjunction with the SPEAK task jointly gave the best portrait of a speaker's fluency. Deleting one of those tasks would only detract from the overall composite picture of a participant's oral fluency.

The aim of the second phase of quantitative analysis of scores was to investigate the impact of task structure (cuing and interactivity) on speakers' display of fluency. It was hypothesized that both the cued and dialogue conditions would provide less communicatively stressful contexts for oral production and speakers' fluency ratings on those tasks would be higher than on the uncued and monologue conditions.
Calculation of the mean ratings of the 46 participants in the monologue and dialogue conditions showed that participants scored higher in the dialogues than in the monologues (dialogue $M = 4.33$, $SD = 0.78$; monologue $M = 3.76$, $SD = 0.98$). However, the mean ratings for the cued and uncued conditions were not very different (cued $M = 3.97$, $SD = 0.91$; uncued $M = 4.12$, $SD = 0.83$). In order to find out whether such differences between means were significant, a repeated measures analysis of variance was conducted in which proficiency level served as the independent variable, cued/uncued and dialogue/monologue conditions respectively labeled “cuing” and “task interactivity” served as factors, and the four fluency task scores served as dependent variables. The results are displayed in Table 2.

Data indicate that, as hypothesized, interactivity had a significant effect on speakers’ display of oral fluency; however, cuing failed to have a significant impact on speakers’ fluency. In addition, there was a significant cuing versus task interaction (see Figure 2). Proficiency level proved to make a significant difference in the scores obtained so that the higher the proficiency level the higher the scores. The same pattern was repeated in the three proficiency levels.

Table 2: Repeated Measures Analysis of Variance

<table>
<thead>
<tr>
<th>Source</th>
<th>$df$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Between subjects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proficiency (P)</td>
<td>2</td>
<td>7.36**</td>
</tr>
<tr>
<td>Within cells</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Within subjects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interactivity (I)</td>
<td>1</td>
<td>39.47***</td>
</tr>
<tr>
<td>$P \times I$</td>
<td>2</td>
<td>0.39</td>
</tr>
<tr>
<td>Cuing (C)</td>
<td>1</td>
<td>2.65</td>
</tr>
<tr>
<td>$P \times C$</td>
<td>2</td>
<td>0.25</td>
</tr>
<tr>
<td>$I \times C$</td>
<td>1</td>
<td>9.71**</td>
</tr>
<tr>
<td>$P \times I \times C$</td>
<td>2</td>
<td>2.25</td>
</tr>
</tbody>
</table>

Note. **p < .01. ***p < .001.

In Figure 2, the interaction between the two task structure variables (cuing and interactivity) can be seen more clearly. In the dialogue condition, absence of cuing led to higher fluency scores, whereas in the monologue condition, the difference in performance was not significantly different when cuing was present. This means that although only one of the task structure variables proved to cause variability in fluency ratings, when both task structure variables were
taken in conjunction (resulting in the four speaking tasks used in the study), they yielded significant variability in raters' assessment of participants' fluency.

**Figure 2: Task Interactivity by Cuing**

![Graph showing task interactivity by cuing](https://example.com/graph.png)

In conclusion, quantitative analyses of the impact of the context of speech production on raters' assessments of speakers' fluency revealed that task interactivity did have an effect. Speakers appeared to be more fluent in dialogues than monologues. The presence of external support for a speaker's discourse—cuing—did not prove to be statistically related to how fluent someone appeared to be. However, when the context of speech production combined different degrees of interactivity with different degrees of cuing, speakers' fluency varied. When participants had to produce a longer stretch of uninterrupted discourse without any external aid—in the uncued monologue—, their fluency decreased. Conversely, in a context of speech production when they interacted with an interlocutor and had cues to guide their performance—the cued dialogue—, their fluency increased. In short, the analyses confirmed the hypothesis that speakers' display of fluency is dependent on the context in which it is elicited.

6. Discourse Analysis: Markers of High and Low Fluency
In order to answer the second research question regarding the differences between fluent and disfluent speech, discourse analyses were performed on the transcriptions of the 24 speech samples produced by six participants: the three least and the three most fluent speakers. The selected case studies of high fluency were Daniela, Roberto, and Eliana. The low fluency speakers were Rubens, Robinson, and Lucia. This section describes four features of speech: amount of speech, speech rate, talk unit length, and fluent unit length (i.e., length of run).

**Amount and Rate of Speech**

Amount of speech was calculated in semantic units; that is, words or comprehensible parts of words. Table 3 presents the mean amount of speech and the speech rate for each individual speaker and the aggregate results for the high and low fluency speakers, considered as fluency groups. In Table 3, the column headed “Amount of speech in su’s” indicates the mean amount of speech participants produced in the four speaking tasks; the column headed “Talk ratio” presents the proportional amount of talk participants engaged in comparison with the researcher who served as the interlocutor in the dialogues. For instance, the high fluency speakers’ mean amount of speech produced was 603 semantic units while the researcher uttered 297 units (not shown in Table 3). Thus, the total amount of speech produced was 900 semantic units, of which the participants produced 68%.

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Amount of speech in su’s</th>
<th>Talk Ratio</th>
<th>Speech Rate in su’s/minute</th>
<th>Task Duration in mins:secs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Fluency Speakers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roberto</td>
<td>793</td>
<td>77%</td>
<td>184</td>
<td>4:56</td>
</tr>
<tr>
<td>Eliana</td>
<td>358</td>
<td>58%</td>
<td>151</td>
<td>3:10</td>
</tr>
<tr>
<td>Daniela</td>
<td>659</td>
<td>67%</td>
<td>165</td>
<td>4:49</td>
</tr>
<tr>
<td><strong>Group Mean</strong></td>
<td>603</td>
<td>68%</td>
<td>166</td>
<td>4:18</td>
</tr>
<tr>
<td><strong>Low Fluency Speakers</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lucia</td>
<td>365</td>
<td>57%</td>
<td>98</td>
<td>4:38</td>
</tr>
<tr>
<td>Rubens</td>
<td>251</td>
<td>60%</td>
<td>78</td>
<td>3:51</td>
</tr>
<tr>
<td>Robinson</td>
<td>314</td>
<td>60%</td>
<td>101</td>
<td>3:59</td>
</tr>
<tr>
<td><strong>Group Mean</strong></td>
<td>310</td>
<td>59%</td>
<td>92</td>
<td>4:10</td>
</tr>
</tbody>
</table>

Note. su = semantic unit.

Results indicate that two of the three high fluency speakers spoke more than the low fluency ones. Roberto and Daniela spoke considerably more than the less fluent participants, but Eliana did not. Furthermore, all high fluency speakers spoke comparatively much faster. Speech rate, thus, turned out to be the more significant discourse feature in discriminating between the two fluency groups. The importance of speech rate had been noted not only in the literature review.
of fluency-related features, but also in the development of the fluency rating scale where speech rate appears as one of the indicators of fluency in the descriptors.

High fluency speakers tended to speak more than their less fluent counterparts not only in monologues but also in dialogues. Thus, the talk ratio varied across with fluency in a pattern in which the more fluent speakers tended to engage in more talk compared with their interlocutor than did the low fluency ones. Roberto, who was considered by the raters as the most fluent speaker of all, best exemplified the pattern that fluent speakers spoke much and fast. Roberto’s speech can be described as very rapidly uttered, sounding "relentlessly fluent" or "hyperfluent" (Butler-Wall 1986) at times. For instance, in the cued monologue, by the time I had uttered my first 100 semantic units, Roberto had already managed to put in almost 600 units! His speech rate was not only faster than mine, but he also struggled for the floor and monopolized it. His and Daniela’s use of overlapping speech as a way of producing of more and faster speech is examined below in Excerpts 2 and 3.

Even when high and low fluency speakers’ amount of speech did not differ appreciably, as was the case in Eliana’s and Lucia’s cued dialogues, speech rate distinguished the very fluent from the very disfluent. In fact, in the role-play with Lucia, I spoke proportionally more than in any dialogue in which I participated. This happened throughout the cued dialogue (Excerpt 1 below), but especially toward the end of the interaction when we still had to make the definite arrangements for the classes. At that point, she either appealed for me to take over (see lines indicated by →) or simply failed to attempt to carry out the instructions specified in the cues. For example, since she had not said when she wanted to have classes or when we should start—that is, the day(s) of the week and time—I finally had to ask her myself.

(1) Lucia’s cued dialogue

```
1 R: [Ah] I see. And when is this person coming?
2 P: When?
3 R: Yeah.
4 P: Uhm: in ( ) December.
5 R: Well. That’s good. At least you have some time to prepare. And he doesn’t speak any
6 Portuguese?
7 P: No. Not at all.
8 R: I see. So, yeah, you definitely need to do some work. So at this point, I mean, it’s kind of
9 late for… to get into the regular course. Are you thinking about the private classes?
10 P: Maybe. I think it’s/… it’s a good idea private… private classes
11 (...)
12→ R: [And you said… Did you say you’d prefer after lunch? Or you didn’t say … I… I don’t
13 remember.
14→ P: [No. I didn’t say] (laugh)
15 R: [Wh… When] would you prefer?
16 P: Uh: in the morning I think it’s better.
17 R: [So… in which case, could you come to my house?
18 P: Yeah. No problem.
```
Lucia’s pattern of reduced speech production is illustrative of the low fluency speakers. Her lack of strategies for fluent speech production made her struggle to communicate so intense that, at one point of the interaction, she almost gave up trying to speak.

On the other hand, fluent speakers seemed to enjoy speaking or to be willing to do so. They optimized their use of time so that they said a lot in a given period of time. Speaking seemed to come to them with ease and, occasionally, they were eager to speak even to the point of competing for the floor with an interlocutor. For instance, one index of fluency, which became apparent in the dialogues, was actual time engaged in speech production. In the dialogues, high fluency speakers engaged in longer periods of actual speech production in comparison with low fluency speakers. One way of accomplishing that was overlapping with me, their interlocutor. Both Roberto and Daniela often overlapped with my speech, making maximum use of the time. For example, in the cued dialogue Daniela spoke 63% of the time. Between the two of us, we used 98% of the total task duration time.

The amount of overlap in Roberto’s cued dialogue in Excerpt 2 shows how he managed to produce so much speech throughout the conversation. The impressive amount of speech rapidly produced and the overlapping allowed him to direct the course of the conversation. By the end of the interaction, I finally acknowledged and gave in to the pressure (see line 5).

(2) Roberto’s cued dialogue

<table>
<thead>
<tr>
<th></th>
<th>R:</th>
<th></th>
<th></th>
<th>P:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[Twenty-five] dollars an hour. So it would be ten days. ( ) So (calculating out loud)... two hundred and fifty dollars and uh: (xxx) think; that’s would be fine. Can you do that?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Yeah.( ) It would be eh: ...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4+</td>
<td>[Can you do that for me?]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>[It’s doable... Not (laugh)... a little (laugh)... to (laugh) pressure but yeah&lt;&gt; we could do that</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
|7  | [Oh. That] would be fine, so we can start immediately believe that... uh Can you... can you come to...tomorrow to:...
9 R: [To the] company. Sure.
10+ P: [C... c... can you come tomorrow “ent” so at seven o’clock a.m., and I will wait still... we’ll
be waiting for you there.
12 R: You will b...you will be there for sure ( ), right?
13+ P: [I... I... I don’t] want... Will you help me because this is gonna be ( ) really:, you know...
really important, so uh:
15 R: [OK]
16 P: [I: ] agree
17 R: [That sounds like a good deal.
18 P: OK. Thank you very much for your attention and see you tomorrow.
19 R: See you then.

Once Roberto had won the battle of getting me to go all the way across Sao Paulo very early
in the morning to give him private classes, the urgent need to communicate abated, and he made
a perfectly uninterrupted turn which did not overlap my speech in line 17.

Daniela’s uncued dialogue (Excerpt 3 below) is similarly representative of her eagerness to
speak and the fast rate of the high fluency speakers. Like Roberto, she occasionally spoke fast to
the point of sounding hyper. This was the case when she got carried away with a joke about
Fernando Collor de Mello, the Brazilian President in office at the time.

(3) Daniela’s uncued dialogue

1 R: [Uhm... And I haven’t been to Brasilia. Would you say that Brasilia is an interesting place to
go?
3 P: [Oh: Brasilia’s] great! There you... you could see: Collor, you know, "Fernando Collor de
Mello" is our... (breath) the best president we... we... we... we: /al/.../al/.../al/.../always/
had.
5 R: [Oh:]
7 P: And: Yeah... He’s... he’s really good; if you go to Brasilia, I can go with you. (breath)
I’ve... I’ve been there for three times only this year; yeah, it was my... my last trip
9 R: But just to go:
10 P: [Just].... just to see Collor because he... he’s a fantastic person; I’m sure you would LOVE
him (laugh)
12 R: [Well] (laugh)
13 (She becomes a little hyper; reference to access ramp to Brazilian White House)
14 P: [Yeah:] (xxx) up and down... up and down that building, you know, “Palacio do Planalto”.
Yeah It’s really good! I think you would love going there! (laugh)
16 R: Well, you see, I have two problems. First of all, I... I don’t really think I can afford to pay for
somebody else to come with me, and the other thing: aren’t you working right now and
18 studying?
19 P: Oh but this is not the problem.
20 R: This is not a problem?
21 P: No. Not a problem. Yeah I could... I could go... I could have some... some “license” from
my work, and uh ( ) well my college is... is... is... ( ) it’s not the... the problem too. So I
could go with you just to show Collor, you know
In contrast, low fluency speakers acknowledged difficulty expressing themselves orally, appeared almost reluctant to speak and, consequently, did not speak as much as their more fluent counterparts (except for Eliana) either when speaking alone or with an interlocutor. In the dialogues, their concern was primarily to find ways of making themselves understood, and this often prevented them from effective conversational interaction.

To summarize, this phase of the discourse analysis has examined two discourse features, which were related to variability in oral fluency: amount and rate of speech. Results showed that amount of speech did not prove to be the major factor in distinguishing between high and low fluency speakers. Although there was a tendency for more fluent speakers to say more than the less fluent, differences in other features such as speech rate were more salient.

The Effect of Task Structure on Amount and Rate of Speech

Profiles of high and low fluency speakers were analyzed across contexts of speech production to uncover patterns of fluency and disfluency that marked and differentiated high and low fluency groups across speaking conditions and tasks. Results are reported in Table 4 and indicate that both high and low fluency groups' amount and rate of speech were affected by the context of speech production. Table 4 presents aggregate data on the impact of task on the amount and rate of speech while Table 5 displays shows the effect on fluency of one the task structure variables: interactivity (i.e., dialogue versus monologue condition). The following patterns are noteworthy in these results: (a) Both high and low fluency groups produced the greatest amount of speech in the uncued dialogue; (b) High fluency speakers spoke the fastest in the cued dialogue; (c) Low fluency speakers spoke fastest in dialogue tasks; (d) High fluency speakers spoke least in the cued monologue and slowest in both monologue tasks; and (d) Low fluency participants spoke least in the uncued monologue and slowest in the cued monologue.
Table 4: Comparison of Amount and Rate of Speech Across Task by High and Low Fluency Speakers

<table>
<thead>
<tr>
<th>Fluency feature</th>
<th>High fluency speakers</th>
<th>Low fluency speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cued monologue</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amount of speech (su’s)</td>
<td>412</td>
<td>298</td>
</tr>
<tr>
<td>Task duration (mins:secs)</td>
<td>2:34</td>
<td>4:17</td>
</tr>
<tr>
<td>Speech rate (su’s/min)</td>
<td>159</td>
<td>68</td>
</tr>
<tr>
<td><strong>Uncued monologue</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amount of speech (su’s)</td>
<td>621</td>
<td>260</td>
</tr>
<tr>
<td>Task duration (mins:secs)</td>
<td>4:02</td>
<td>3:00</td>
</tr>
<tr>
<td>Speech rate (su’s/min)</td>
<td>152</td>
<td>86</td>
</tr>
<tr>
<td><strong>Cued dialogue</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P’s amount of speech (su’s)</td>
<td>575</td>
<td>329</td>
</tr>
<tr>
<td>R’s amount of speech (su’s)</td>
<td>292</td>
<td>283</td>
</tr>
<tr>
<td>Talk ratio</td>
<td>66%</td>
<td>54%</td>
</tr>
<tr>
<td>Task duration (mins:secs)</td>
<td>4:46</td>
<td>4:49</td>
</tr>
<tr>
<td>P’s speech rate (su’s/min)</td>
<td>181</td>
<td>107</td>
</tr>
<tr>
<td>R’s speech rate (su’s/min)</td>
<td>204</td>
<td>221</td>
</tr>
<tr>
<td><strong>Uncued dialogue</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P’s amount of speech (su’s)</td>
<td>805</td>
<td>351</td>
</tr>
<tr>
<td>R’s amount of speech (su’s)</td>
<td>302</td>
<td>225</td>
</tr>
<tr>
<td>Talk ratio</td>
<td>73%</td>
<td>61%</td>
</tr>
<tr>
<td>Task duration (mins:secs)</td>
<td>5:51</td>
<td>4:32</td>
</tr>
<tr>
<td>P’s speech rate (su’s/min)</td>
<td>173</td>
<td>108</td>
</tr>
<tr>
<td>R’s speech rate (su’s/min)</td>
<td>229</td>
<td>223</td>
</tr>
</tbody>
</table>

Note. There were 3 speakers in each fluency group. su = semantic unit. su’s/min = semantic units per minute. P = participant. R = researcher.

An examination of the data in Table 5 reveals that task interactivity affected the amounts produced by the two fluency groups. Both high and low fluency participants spoke more in the dialogues than in the monologues. However, across task structure conditions, with the exception of Eliana, the very fluent spoke more in both absolute and relative terms.

Cuing helped tease apart the fluent from the disfluent speakers: Both groups responded to its presence or absence in clearly distinct ways. Clearly, cuing had a large impact on the amount of speech produced by the very fluent but a negligible impact on how much the low fluency participants spoke. The fact that the highly fluent spoke much more in the uncued tasks, as anticipated, may mean that contexts of more contrived speech production (i.e., cued tasks) did not allow for discourse differences between speakers to become fully visible. Yet, when the task gave participants more flexibility and freedom to display their fluency (i.e., in the uncued tasks), profiles of fluent and disfluent language use became more apparent.
Table 5: Effect of Interactivity on Speakers’ Amount and Rate of Speech

<table>
<thead>
<tr>
<th>Fluency feature</th>
<th>High fluency speakers</th>
<th>Low fluency speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount of speech (su’s)</td>
<td>517</td>
<td>279</td>
</tr>
<tr>
<td>Task duration (mins:secs)</td>
<td>3:18</td>
<td>3:38</td>
</tr>
<tr>
<td>Speech rate (su’s/min)</td>
<td>159</td>
<td>77</td>
</tr>
<tr>
<td>P’s amount of speech (su’s)</td>
<td>690</td>
<td>340</td>
</tr>
<tr>
<td>R’s amount of speech (su’s)</td>
<td>297</td>
<td>254</td>
</tr>
<tr>
<td>Talk ratio</td>
<td>70%</td>
<td>57%</td>
</tr>
<tr>
<td>Task duration (mins:secs)</td>
<td>5:10</td>
<td>4:40</td>
</tr>
<tr>
<td>P’s speech rate (su’s/min)</td>
<td>177</td>
<td>108</td>
</tr>
<tr>
<td>R’s speech rate (su’s/min)</td>
<td>217</td>
<td>221</td>
</tr>
</tbody>
</table>

Note. There were 3 speakers in each fluency group. su = semantic unit. su’s/min = semantic units per minute. P = participant. R = researcher.

Results also indicate that task interactivity affected the speech rate of the two fluency groups: Both high and low fluency speakers spoke faster in the dialogues than in the monologues. However, the gap in speech rate between the cued and uncued conditions was much wider in the low fluency speakers’ monologues. Table 5 shows that cuing affected the speech rate of the very fluent and the very disfluent in opposite ways. Cuing caused the high fluency participants to speak slightly faster but it caused the low fluency speakers to speak more slowly.

For instance, Lucia’s worst performance, the cued monologue, was the task in which she spoke the least and the most slowly. In her feedback, she acknowledged the tremendous difficulty she experienced carrying out the task. Nonetheless, she said that the cued monologue “best put [her] English to test” and best reflected her spoken English ability. The raters noted her conspicuous difficulty evidenced by a slow rate, a factor that contributed to her coming across as a very disfluent speaker. The text of her cued monologue is presented in its entirety in Excerpt 4 to provide an idea of how she plowed through the plot. At one point (see line 15), she was so frustrated that she asked for “time out” in Portuguese: “Pera!”.

(4) Lucia’s cued monologue

1 My lover Jack and I saw in the news that it was a /jewry/... there was a /jewr/... /jewerly/... /jewerly/ uh: on sale on: fifth uh uh I s... I mean uh: on fifth... The number the... the... the /store/ was fifteen thirty-one. So ( ) we:<...: we went in the /store/, and we saw the... ( ) the... ( ) we saw the...the... the address was right, and ( ) uh: I was looking for a... a ring, and... ( ) And: we entered in the /store/, and we talk with the salesman. Uh: he offered uh a ring for twenty dollars.
2 I... I... ( ) I was surprised because the... the ring, in fact, uh: th... the price... the... the original price was one thousand dollar. This is a big sale; th... th... this was a big sale I... I think, and...
The salesman said the ring was not a fake; what’s was real, and my Jack gave two dollars (laugh) Let me let me explain
(laugh) again uh gave twenty dollars for a salesman, and everybody was happy. But when
when we... we realize... when we we /looked/ at ring uh: “/with more precision”, we realize that the ring it was fake. The... an expert said that to... to us, and... so we... we resolve... we (sigh) we resolve to come back to the /store/ to complain to... (breath) The... the ring was fake, and ( ) we: wanted to... ( ) to/... to “pick” my uh... ( ) our twenty dol... uh uh ( )/two/ hundred dollars. And...and the salesman uh... This is (sigh) “Pera!” ( ) And the... ( ) the salesman uh: said in... ( ) at the same time ( ) uh: “I’m going to call the police!” At the same time with Jack, so the... the... the two of ( ) uh:... The Jack and the salesman said: “Go ahead” at the same time... at the same time too, ( ) and we laugh... we /laugh/ because this is a... th... this is... this was a... a coincidence, and we realize uh: that ( ) uh the salesman and Jack and I was lying uh because ( ) we//... we were a kind of /thief/ bo... both. And ( ) uh: ( ) all with...
( ) all almost... ( ) almost the... the... the... the whole situation finished in pizza (laugh).

It had been hypothesized that all speakers would benefit from cuing, not only the highly fluent. One interpretation of this differential impact of cuing on the two fluency groups’ speech rate is that the cues posed constraints on speech production and cornered speakers into saying certain things or exhibiting knowledge of a particular set of words, expressions and communicative functions, which worked against the low fluency speakers. To summarize, the evidence provided by the analysis of amount of speech and speech rate in the discourse of the six case studies indicates that high fluency speakers tended to speak more and much faster than the low fluency ones. Speech rate, in particular, was a feature of fluent speech that discriminated between the two fluency groups.

Across fluency groups, the context of speech production affected discourse variability and fluency ratings. Task interactivity affected the amount of speech produced: All speakers spoke more in the dialogues than in the monologues. Cuing affected the fluency groups differently with respect to amount of speech. Whereas it had negligible impact on how much the low fluency speakers spoke, it made a noticeable difference for the highly fluent participants who spoke much more in the uncued condition. However, when both Interactivity and Cuing were considered jointly, both fluency groups spoke the most in the uncued dialogue.

The context of speech production also affected participants’ speech rate. Task interactivity affected the two fluency groups similarly: Both had higher speech rate in the dialogues than in the monologues. Cuing had a differential impact on the two fluency groups’ speech rate: in the cued condition, when compared to each group’s performance in the uncued condition, high fluency speakers spoke faster while low fluency participants spoke more slowly.

In addition, the ratings of the low fluency speakers—but not of the high fluency participants—were consistently influenced by task interactivity: Low fluency speakers’ ratings were higher in the dialogues than in the monologues. Lastly, all speakers sounded more fluent in the uncued dialogues than in the cued dialogues.

*Talk Unit and Fluent Unit Length*
Besides amount of speech and speech rate, the other salient discourse features that marked the difference between the very fluent and the very disfluent were talk unit and fluent unit length, which were calculated in the monologue tasks. A talk unit was operationalized as any continuous stretch of two or more semantic units. The markers of talk unit boundaries were presence of a silent pause of one second or more and “blocks” or “abrupt halting of speech” (Postma, Kolk, and Povel 1990:19). Neither micropauses (unfilled pauses smaller than one second) nor filled pauses were counted as talk unit boundaries. Table 6 presents data on speakers’ talk unit length, including the longest talk unit uttered and the most frequent talk unit length (i.e., the modal length) with its frequency of occurrence in each speaker’s monologic discourse. Modal figures have been presented instead of means because the latter were artificially inflated by the high values of a few outliers in the upper part of the distribution (in the case of the high fluency participants).

Table 6: Speakers’ Talk Unit Length in Monologic Discourse

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Longest talk unit</th>
<th>Modal length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cued monologue</td>
<td>Uncued monologue</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High fluency speakers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roberto*</td>
<td>65</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eliana</td>
<td>37</td>
<td>56</td>
</tr>
<tr>
<td>Daniela</td>
<td>57</td>
<td>60</td>
</tr>
<tr>
<td>Group mean</td>
<td>53</td>
<td>60</td>
</tr>
</tbody>
</table>

| Low fluency speakers |                   |               |               |                |
| Lucia*              | 16                | 26            | 4.5 (49%)     | 4.5 (64%)     |
| Rubens              | 16                | 34            | 4.5 (65%)     | 4.5 (65%)     |
| Robinson            | 27                | 39            | 4.5 (51%)     | 4.5 (65%)     |
| Group mean          | 20                | 33            | 8.2           | 4.5           |

Note. All measurements are in semantic units (su’s). Frequencies measured by interval; mid-interval points shown. * Bimodal distribution

The extremes of the talk unit distribution, namely 2 and 68 semantic units, were the longest and shortest talk unit length values that emerged from the data. Six-point length intervals were devised, and the most frequent mid-interval point for each speaker calculated. In Roberto’s and Lucia’s cued monologues, there are two mid-interval points because their distributions were bimodal; that is, there were two mid-interval lengths that occurred with identical frequency. Because there was no variability in length of the shortest talk units, such figures are not reported.

The figures clearly indicate that low fluency speakers uttered much shorter talk units than the fluent speakers. This was an important discourse feature that contributed to the low fluency.
participants’ speech being perceived as more fragmentary, discontinuous and/or choppy. This finding has to be viewed with caution lest it become a simplistic and erroneous attribution of fluency to talk unit length. In Excerpt 5, for instance, talk unit length alone was not an indicator of fluency since the units varied in the quality of the speech produced. This excerpt shows Rubens’ talk units number 1 and number 8, with respectively 34 and 26 semantic units. These were fairly long talk units for Rubens (since his modal length ranged from 6 to 8 semantic units), but whereas the first was quite smooth and fluent the latter was the exact opposite. The impression of choppiness in talk unit 8 stemmed from the fact that words were uttered a few at a time, not connected as a fluent stretch. Thus, the quality of the speech within the talk units, rather than their length, accounted for the fluency and/or disfluency.

(5) Rubens’ uncued monologue

Talk unit 1
{"Entao”. [I’d like to... to talk about the...][some problems] eh: [that there is...][there are] <> in//...[in Sao Paulo because Sao Paulo is a big city ], and:... <> eh then [there are big problems too].}

Talk unit 8
{I:... <> [I talk... ] <> [I’d like to talk about...][about the:...] [the...the problem] <> [that I have] for: <> two:... <> [two months...][two months after] eh:... before <> and:...}

In sum, both within individuals and across fluency groups talk unit length was an important index of fluency but, by itself, failed to reveal more subtle differences in the speakers’ discourse. For example, segments of identical talk unit length within the discourse of the same speaker performing one task (e.g., Rubens’ uncued monologue) varied qualitatively with respect to smoothness of production and perceived fluency.

The fluent unit (Pawley and Syder 1983) or length of run (Olynyk et al. 1990) is defined as continued speech without any pause or hesitation (pp. 142-143). In the present study, a fluent unit was operationalized as any continuous stretch of running speech consisting of two or more semantic units literally uttered in one breath or nearly so. Consequently, fluent units occurred within talk units. Table 7 shows that the mean fluent unit length was greater in the speech of the very fluent and that increased fluent unit length was directly proportional to increased fluency, so the longer the fluent units, the more fluent the speech and vice versa.

Eliana’s mean fluent unit length was the lowest of her fluency group, confirming her previous pattern of appearing to be fluent because of her strategic competence rather than on the merits of her linguistically derived fluency. Lucia, who is a very articulate speaker in her native language, had the highest mean fluent unit length of the low fluency speakers. She sounded more fluent than Rubens but not Robinson; the reason being that Lucia’s strategic failure to maintain an air of fluency undermined the linguistically derived discourse feature of greater mean fluent unit length that would have otherwise accounted for increased oral fluency.

In order to address the question of how fluent unit length affected fluency ratings, the data were compared with the speakers’ ranked fluency ratings. As anticipated, there was clear
evidence that fluent unit length was directly proportional to the fluency ratings speakers from both fluency groups received on their performance on the monologue tasks. For example, in the high fluency group, Roberto scored lower in the uncued than in the cued monologue. Coincidentally, his mean fluent unit length was also smaller when his fluency rating was lower. For both Daniela and Eliana, the negligible difference in mean fluent unit length (i.e., smaller than or equal to .5 semantic units) between each speaker’s two monologues was paralleled by, respectively, identical and almost identical scores on the two monologues.

Table 7: Speakers’ Fluent Unit Length in Monologic Discourse

<table>
<thead>
<tr>
<th>Speaker</th>
<th>Longest fluent unit</th>
<th>Mean fluent unit length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cued monologue</td>
<td>Uncued monologue</td>
</tr>
<tr>
<td><strong>High fluency speakers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roberto</td>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td>Eliana</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>Daniela</td>
<td>20</td>
<td>24</td>
</tr>
<tr>
<td><strong>Group mean</strong></td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td><strong>Low fluency speakers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lucia</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>Rubens</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Robinson</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td><strong>Group mean</strong></td>
<td>9</td>
<td>11</td>
</tr>
</tbody>
</table>

*Note.* All measurements are in semantic units.

The low fluency group confirmed the pattern of the highly fluent speakers. Lucia’s higher mean fluent unit length was attained in the cued monologue which also had a higher rating than the uncued monologue. Rubens’ and Robinson’s negligible difference in mean fluent unit length between each participant’s two monologues was congruent with each speaker’s identical scores on both monologue tasks.

To summarize this section, mean fluent unit length was a discourse feature that discriminated between the two fluency groups, helping to account for their different fluency ratings. Furthermore, even within a single speaker’s discourse, variability in fluent unit length coincided with episodes of greater or lesser fluency. For instance, in Excerpt 6, Roberto’s talk unit 11, which contained much shorter fluent units than the preceding talk unit, sounded much less fluent.
(6) Roberto’s uncued monologue

Talk unit 10  { [And, you know], [here the situation is that] <-> [it’s a big… big city], [and this city is] uh: <-> [aggressive in a certain way, so [when you come in] (breath) [be careful the moment you leave the airport…] [the moment you leave your… your plane].}

Talk unit 11  {[But even so] [don’t be afraid ’cause after some..] some /months/ (breath) you <-> [will feel like a citizen], and you’ll… [you’ll cope with the problems we have here], <-> [and you’ll learn] [how to enjoy] [living in a big city] <-> with… [with the help of all…] [all your “paulista” friends.]

At the end of talk unit 11, to culminate the segment of reduced fluency, Roberto stopped speaking and took a fairly long pause. It surfaced later in the oral interview that he was unsure whether to continue speaking or to stop at that juncture. This hesitation detracted from his fluency, and the shorter fluent units signaled that.

As to the effect of task structure on speakers’ talk unit length in the monologue condition, cuing only affected the low fluency group. These speakers had higher talk unit modal length when cuing was present, which might be interpreted as their response to the potentially facilitating role of cuing. For the high fluency speakers, however, there was no difference in their talk unit length in the cued and uncued monologues.

When the totals from the high and low fluency group were analyzed, cuing seemed to have some impact on speakers’ mean fluent unit length. For the high fluency participants, the mean length was somewhat higher in the cued monologue than in the uncued monologue. For the low fluency speakers, on the other hand, the results were the opposite. A closer scrutiny of the data, case by case, revealed that cuing had either no impact or a negligible impact on fluent unit length in the speech of Eliana, Daniela, Rubens and Robinson. Furthermore, it affected Roberto’s and Lucia’s discourse in opposite ways. When it was present (i.e., in the cued task), Roberto’s fluent units were longer and Lucia’s were shorter than when it was absent (i.e., in the uncued monologue).

7. Summary of Findings

As hypothesized, task structure had an impact on speakers’ fluency ratings and discourse variability. The statistical analyses of the entire sample (N = 46) showed that the interactivity present in the context influenced speakers’ display of fluency. Speakers were more fluent when performing the dialogues and came across as less fluent when performing the monologues. Although cuing per se was not statistically related to how fluent someone appeared to be, the interaction of task interactivity (dialogue/monologue) and cuing (cued/uncued) affected speakers’ fluency ratings. Thus, the analyses revealed that speakers’ display of fluency was context dependent when interactivity and cuing were considered jointly. For instance, in a context of speech production when they interacted with an interlocutor and had cues to guide
their performance, as was the case in the cued dialogue, their oral fluency improved. Yet, when they produced a longer stretch of uninterrupted discourse without any external aid as was the case in the uncued monologue, they sounded more disfluent.

The analyses of the speech of the three most fluent and the three least fluent speakers showed the nature of their discourse differences from the psycholinguistic perspective of speech production and from the sociolinguistic angle of context-related fluency. One pattern that emerged was that speech rate is a feature of fluent speech that distinguishes the two fluency groups. The more fluent speakers spoke much faster than their less fluent peers. Moreover, high fluency speakers tended to speak more than the low fluency ones.

Speech rate, amount of speech, and fluency ratings in the case studies of the six speakers were context dependent. Task interactivity affected all six similarly: They spoke faster in the dialogues than in the monologues. Cuing had a differential impact on speakers' speech rate: In the cued condition, the speech rate of high fluency speakers increased whereas the others decreased. Task interactivity had an effect on amount of speech produced. Contrary to my previous expectations, all speakers spoke more in the dialogues than in the monologues.

The impact of cuing was felt differently by the two fluency groups. Whereas it had negligible impact on how much the low fluency speakers spoke, it made a noticeable difference for the others, who spoke much more in the uncued condition. The presumably facilitating role of cuing in speech production failed to occur either at the level of the entire sample or in the individual analyses. Although cuing generated variability in speakers' discourse, the qualitative differences did not affect speakers' fluency ratings. This was not the case when the effect of task interactivity was analyzed. As anticipated, the statistical analysis showed that the 46 participants' ratings were higher in the dialogue than in the monologue. For the selected six speakers, fluency ratings were higher in the uncued dialogue than in the cued dialogue but did not differ between the two monologues.

Talk unit and fluent unit length were directly proportional to oral fluency so that the longer the talk units and the fluent units, the more continuous, that is, fluent the speech. Thus, highly fluent speakers produced longer talk units and fluent units than the low fluency speakers. Of the two discourse features, fluent unit length was found to be the more accurate index of oral fluency, better reflecting the degree of continuity and flow of participants' speech.

As to the impact of the context of speech production on speakers' talk unit and fluent unit length in the monologues, only one pattern marked the difference between high and low fluency groups: Low fluency speakers had higher talk unit length when cuing was present (in the cued monologue). When external support to organizing their discourse was present, these speakers were able to produce longer stretches of speech. High fluency speakers' talk unit and fluent unit lengths, on the other hand, were not affected by the context of speech production. Lastly, whereas talk unit length was not systematically related to speakers' fluency ratings, fluent unit length was a strong indicator of variability in speakers' fluency ratings. In effect, the longer a speaker's mean fluent unit length, the higher his/her fluency ratings and vice versa.

8. Conclusions and Implications
Only by examining the effect of context on discourse production as it emerges from task type, participant roles, and speakers’ perceptions of task demands (Samuda 1990), will we be able to establish a theory of fluency that incorporates situational variables and is sensitive to the ways in which fluency is differentially displayed across task types and settings. Once we obtain a model of variable fluent speech production across tasks, then we may control for task effect in order to examine the fluency-related features that comprise fluency and how they are related to the domains of linguistic, discourse, sociolinguistic and strategic competence (Canale 1983; Canale and Swain 1980).

**Implications for a Theory of Fluency and Proficiency**

The findings of this study indicate that fluency, from the psycholinguistic perspective of the speaker’s planning and production, can be defined as automaticity of speech delivery. To an interlocutor or rater, fluency is defined as the speaker’s ease of expression judged from the perceptions of the speaker’s rate of articulation and ability to maintain an appropriate speech flow according to contextual demands. From a sociolinguistic perspective, fluency is context dependent and may vary within the same individual across tasks.

As reported in Ejzenberg (in press), fluency is a feature of language proficiency that is made possible by several of its components but mostly by the speaker’s strategic competence. Analogous to public speaking, fluent speech has to do with projecting an image of being fluent, maintaining an “air” of fluency and managing shortcomings in linguistic and discourse competence. Discourse competence, manifested in the speaking mode as the ability to construct and organize extended stretches of speech, was found to be a component of proficiency with great bearing on oral fluency.

**Implications for Instructional and Testing Practices**

Task-based research in second language, of which this study is representative, has also proposed a paradigmatic change in curricular theory and practice that rejects synthetic, reductionist syllabi (Long and Crookes 1992 drawing on Wilkins 1974) that primarily focus on what is to be learned (Long and Crookes 1992:29). This approach to curriculum design has been criticized for segmenting language into decontextualized and unauthentic parts to be pieced together—synthesized—by the language learner. In an alternative framework—task-based syllabus design—the focus shifts toward contextualized, process-oriented learning centered around the task as the curricular unit, “the task receives more support in second language acquisition research as a viable unit around which to organize language teaching and learning opportunities” (Long and Crookes 1992:27).

Students should be exposed to tasks of variable structuring in order to be empowered as speakers able to perform in different communicative and testing contexts. As Brown and Yule (1983:19) put it,
the ability to construct ... long turns appears to vary with individuals, in part, no doubt, depending on the opportunity they have had to produce long turns which other people bother to listen to. The ability to produce long transactional turns, in which clear information is transferred, is, we claim, not an ability which is automatically acquired by all native speakers of a language. It is an ability which appears to need adequate models, adequate practice and feedback.

The two authors continue the argument by saying that if for native speakers the difficulty is such, the problem is compounded for the nonnative.

Our greater understanding about testing at higher levels of speaking proficiency should have an impact on testing procedures adopted so that they allow for a display of fluency. The present study suggests that it is important for the field to move toward multitask and portfolio assessment. Such kinds of assessment would solve the problems of the variability in fluency ratings resulting from the use of different tasks.

Above all, this study provided strong evidence to refute conceptualizations of nonnative proficiency and fluency that are grounded on absolute and immutable terms derived by comparison to an unreachable "nativelike" or "near-native" ideal. Corroborating the work of Riggenbach (1989) and other researchers, the study showed that nonnative oral fluency comes and goes according to the context of speech production and the speakers' perceptions of task demands. It is only fair to give nonnative speakers the opportunity to display the different facets of their oral proficiency and fluency. Furthermore, testing instruments designed to assess nonnative fluency should undergo rigorous scrutiny to do away with fluency scales and descriptors that confound proficiency and fluency and/or compare nonnative fluency to native standards.

**Implications for Research**

There are some limitations to the present study in terms of the generalizability of the findings. This study should be replicated in different settings and with different linguistic communities to find out whether the patterns found here are indeed generalizable. Moreover, it is important to analyze the discourse of speakers who are neither very fluent nor very disfluent, to gain more insight into the paths that lead to greater fluency.

This research is of particular significance due to the potential teaching implications. Of special relevance for the field is the need to replicate the study controlling for topic variability to be better able to isolate the effects of cuing and of topic preference on speech production.

Further research should also address the pending question of how task interactivity affects formulaic speech and fluency. Despite the fact that talk or speech units in monologue tasks such as story telling are expected to be longer (Goodwin 1990), what remains to be investigated is the issue that because of their higher level of "communicative stress" (Brown and Yule 1983), monologue tasks may not necessarily yield longer fluent units or more speech than dialogues. New knowledge in this area may provide a firmer foundation for large-scale decisions about the tasks currently used in oral language tests such as the Oral Proficiency Interview and the Test of Spoken English.
Continuing research into the nature of language proficiency (Shaw 1992), accompanied by the questioning of accepted models in the field, needs to generate novel theorizing that is grounded in more research along the conceptual lines of studies like the one reported here. This study clearly highlights the need to disentangle notions of nonnative fluency and proficiency.
Notes
1. The four fluency task prompts were as follows:

Cued monologue: “The Truth Story”, extracted from Shannon (1989). “Who can you trust?” (pp. 148–149) is an eight-frame cartoon series in black and white, five of which have short verbal captions in English (frames 2 and 4–7). The cartoons show a couple (a man and his female companion) who go to a jewelry store that is having a sale (frame 1). After the salesman reassures them that the diamond is real (frame 2), they buy the diamond ring which had been marked down from $1,000 to $2,000 (frame 3). They go to another jeweler who checks the ring with a magnifying glass and says it is fake (frame 4). Meanwhile, the salesman examines the bills he received from the couple and finds out they are counterfeit (frame 5). The couple returns to the jewelers to complain about the fake diamond ring. Both the man and the salesman say they are going to call the police (frame 6), and they tell each other to “Go ahead” (frame 7). Then both start laughing (frame 8).

Prompt for the cued monologue: “This is a sequence of pictures that tell a story. It starts in picture 1 and ends in picture 8. First of all, take a look at this picture sequence and see if you can understand the story. (Pause.) Then I’ll ask you to imagine that this story happened to you last month and to tell the story from that perspective. When you tell the story make sure to tell it in the past and to speak as much as you can. Use your imagination to tell it in as much detail as possible. Feel free to ask me any questions before you start telling the story, but once you start do not stop to ask any questions for any reason. Let me know when you are ready to start.

Uncued Monologue, The Realities of Sao Paulo: A foreign colleague/business partner/friend is coming to Sao Paulo for a couple of months. Since it will be his/her first time here you want to prepare him for some of the realities of living in a large city. Instead of calling him/her up or writing a letter you decide to send him/her a tape to caution him/her about some of the realities of a big city like Sao Paulo. Talk about one event or story that has happened to you or to someone you know to illustrate your point.

Cued Dialogue, Coping with a Presentation: If you work at a multinational company in Sao Paulo, imagine that you will need to give a presentation in English to the company’s General Manager who will be coming from the US to visit the Brazilian branch. (Otherwise, imagine that a foreign colleague/expert will be coming from the US for a short visit to learn about what is going on in your field/line of work in Brazil, and you will need to give him/her a short presentation.) You are feeling apprehensive about the presentation, so you decide to get some help. You will approach one of the school teachers to inquire about private classes. Follow the cues:
(a) Greet teacher and tell him/her you need to talk to her;
(b) Explain what the problem is. Give a brief overview of what you do and what the presentation will be about;
(c) Inquire about the teacher’s availability as well as price, duration and site of private classes;
(d) Ask for an estimate of the frequency and duration of classes needed for you to be prepared for your presentation; and
(e) Take into account the expenses and time involved and come to a decision.

Uncued dialogue, Travel Plans: An American/foreign friend will be in Brazil for a limited period of time and wants to discuss with you about some possible travel plans. Give her as much information and advice as possible.

The comparative fluency measures were as follows:


Cued Monologue in Portuguese: “The Pay Phone story,” from Ligon and Tannenbaum (1990:67), is a ten-frame cartoon series in black and white, only one frame of which has a caption. The story was visually adapted to Portuguese. For instance, the word “cashier” was replaced by “caixa” and “Orange street” was replaced by “Rua Conselheiro Crispiniano”, a street in downtown Sao Paulo). Frame 1 shows a man thinking about finding a telephone. He goes to PG Department Store (frame 2), but the cashier does not let him use their phone (frame 3). He leaves the store and finds a pay phone (frame 4). He goes inside the booth, leaving his briefcase outside (frame 5). He takes a quarter (frame 6), slides it in the phone slot (frame 7) and dials the number (frame 8). While he is animatedly talking
on the phone, with his back turned away from the briefcase, a thief approaches it (frame 9) and sneaks away with it (frame 10).

2. The transcription system is my adaptation of the systems used in Butler-Wall (1986) and Riggenbach (1989) who drew on Jefferson (1972). Transcription conventions: P = Participant, R = Researcher, ... = unfinished or interrupted word or utterance, [ ] = overlap, ( ) = pause, < > = micropause, // = abrupt pause, : = elongation, (xxx) = incomprehensible, and " " = transfer from Portuguese.

3. { } = talk unit and [ ] = fluent unit.
References


I. DOCUMENT IDENTIFICATION

Title: The Role of Task Structure in Oral Fluency Assessment
Author(s): ROSELI EISENBERG
Date: 11/7/97

II. REPRODUCTION RELEASE

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, or electronic/optical media, and are sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document. If reproduction release is granted, one of the following notices is affixed to the document.

If permission to reproduce the identified document, please CHECK ONE of the options below and sign the release on the other side.

- Permitting microfiche reproduction in (4" x 6" film)
- Permitting paper copy, electronic, and optical media reproduction (Level 1)
- Permitting reproduction in other than paper copy (Level 2)

Documents will be processed as indicated, provided quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.
I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce this document as indicated on the other side. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Angelina Epstein
Printed Name: Angelina Epstein
Organization: Montgomery College
Position: Associate Professor
Address: 51 Mannsback Street
Rockville, MD
Tel. No: (301) 251-7373 Zip Code: 20850
Email: REESENBERGMC.CC.MD.Us

III. DOCUMENT AVAILABILITY INFORMATION
(Non-ERIC Source)
If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents which cannot be made available through EDRS).

Publisher/Distributor: ____________________________

Address: ______________________________________

Price Per Copy: __________________________

Quantity Price: __________________________

IV. REFERRAL TO COPYRIGHT/ REPRODUCTION RIGHTS HOLDER
If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

______________________________________________