

New Findings on Fluency Measures across Three Different Learning Contexts

Lorenzo García-Amaya
Indiana University

1. Introduction

According to Schmidt (1992), ‘fluency’ is understood as a primarily temporal phenomenon (i.e., not a vague notion of proficiency, but the way speech is processed and articulated in real time). In this respect, fluency studies in second language (L2) have employed methods previously used in psycholinguistic investigations of the L1, especially in the tradition of Goldman Eisler’s *pausological* studies, which reported that the apparent fluency of native speakers is not free of hesitations. In this sense, second language acquisition (SLA) investigations have also analyzed fluency in terms of temporal variables: words per minute (Freed, Segalowitz, & Dewey, 2004; Lennon, 1990; Segalowitz & Freed, 2004), words per second (Binnenporte, Van Bael, den Os, & Boves, 2005), and syllables per second (Temple, 1992), in combination with the production of “hesitation phenomena” such unfilled and filled pauses (Freed et al., 2004; Lennon, 1990; Segalowitz & Freed, 2004; and Temple 1992).

Surprisingly, previous fluency studies in SLA literature treat speech rate as a rather definite measure and do not account for the existing range that learners show in their speech. Freed (1995) analyzed 45 second segments of speech production, and although more recent studies have increased this measure to 4 minute intervals in order to account for the overall fluency production of the participants in their studies (Freed et al., 2004; Segalowitz & Freed, 2004), a more exhaustive analysis of learner oral production (i.e., of multiple and longer turns) could provide a more reliable account of fluency measures, especially those related to temporal variables and hesitation phenomena.

Previous investigations have examined the use of various dimensions of fluency in three different learning contexts: formal contexts or *at home* (AH), study abroad (SA), intensive domestic immersion (IM). To my knowledge, there is no single study that compares intensive *overseas* immersion programs to other more traditional learning contexts. The benefits of SA programs have been reported in the SLA literature and they include the opportunities for interaction with NSs out of the classroom setting. In fact, there are different types of study abroad programs and it is important to investigate the gains obtained by learners studying abroad in different contexts. The intensive overseas IM program group under examination in the current study is an especially unique group in which participants are committed to a no-English rule during the entire 2 months experience.¹ These rules and the insensitivity of the program makes it different from other study abroad programs. Moreover, the status of such rules in study abroad contexts and the effect they have on measures such as fluency is unclear.

In this paper I argue that previous research on fluency measures does not account for the range of variability that is attested within learners’ oral production. The current study reports on the developmental ranges that exist across 5 levels of proficiency. It also compares obtained in a college-based learning context or AH, with a college based SA environment, and with a high school based intensive overseas IM program. The present paper is organized as follows: the previous literature includes an overview on fluency studies in L2, temporal variables and hesitation phenomena, and different learning contexts. Next, the current study is presented, with research questions, participants,

¹ The abbreviation IM will be used for both the domestic immersion program (e.g., Segalowitz & Freed, 2004) and the overseas immersion program analyzed in the current study. To avoid confusion the adjectives *domestic* or *overseas* will be added before the abbreviation.

data collection, data analysis, results and discussion. Finally, conclusions are drawn, and future directions of research will be presented.

2. Literature Review

Studies of second language (L2) fluency have followed two general trends: first, research concerned with the measurement of quantifiable features; and second, research that has focused on factors that affect the evaluation of learner fluency by the listener. A series of studies conducted in the 1980's investigated the characteristics of fluency in oral production with a special emphasis on hesitations (i.e., Chafe, 1980; Dechert, 1980). Chafe's view of fluency is related to the number of hesitations produced by a speaker. In his pioneering study on "hesitations", he refers to them as "attributable to the speaker's need to find the next focus" (1980: 171). Chafe interviewed twenty adult female speakers of English to recount a movie they had previously watched. He quantified the "hesitation pauses" in the retells and reported that the participants had a larger tendency to hesitate between sentences than within sentences. Dechert (1980) studied pauses and intonation in second language in a case study of George, a learner that had never studied abroad. George had to retell a story before and after spending 3 months in an English speaking country. The results indicated that the first time he retold this fragment he needed 37 words (with 13.3 seconds of pausing time), while only 25 words (with 4.0 seconds of pausing time) the second time. This "improvement" led Dechert to claim that George was more fluent and less hesitant after 3 months in a foreign environment.

Raupach (1980) identified factors such as speech rate, length and position of filled and unfilled pauses, and the frequency of repetitions and self-corrections as measurable indications of L2 fluency. Raupach focused on the cross-linguistic aspect of temporal variables. In his study, 5 French and German undergraduate students gave a description of a cartoon without any text in their native languages. Next, they repeated the performance in their L2 (e.g., German and French respectively). Raupach analyzed the speech rate (number of syllables per minute) of the NSs, which included pause time, and concluded that although the participants were given a text in the cartoon when they had to retell the story in their L2, and therefore the "semantic planning" had already been taken care of for them, pauses occurred more often in L2 than in L1 performance.

A number of studies have also focused on how fluency is perceived at the utterance and discourse level (Olynik, d'Anglejan, & Sankoff, 1990). These authors along with Schmidt (1992) claimed that both the speaker and hearer's perspective of fluency are related. Indeed, several works have studied the relationship between NS judgments' of fluency and measurable characteristics in oral production (e.g., Lennon 1990; Riggensbach, 1991; Temple, 1992). Lennon reported that fewer unfilled pauses per "t-unit", and fewer "t-units" followed by a pause were statistically significant indicators of fluency development during a study-abroad experience.² Riggensbach also analyzed excerpts of audiotaped dialogues of six Chinese learners of English; NS judges deemed 3 of them as "highly fluent" while the other 3 were rated as "highly non-fluent". These findings indicated that a faster rate of speech and fewer unfilled pauses contribute to a more fluent perception of speech production. Temple (1992) took 42 speech samples, 20 of approximately one minute long extracts of interviews of French natives, taken from a commercial cassette, and the other 22 samples from spontaneous spoken French of approximately 2 minutes length produced by eleven second year university students (two times with the same students: 11 at the beginning and 11 at the end of the semester). She reported that learners had "greater demand on time" (p. 42), which was due "to lack of automatization and need to use controlled processes" (p. 42).

A number of fluency studies have also focused on the impact that context of learning has on this facet of SLA (DeKeyser, 1991; Freed, 1990, 1993).³ Following this approach, Segalowitz and Freed

² The term T-unit was coined by Hunt in 1965, who defined this concept as "one main clause with all subordinate clauses attached to it" (p. 20). Although this measure has been widely used for the analysis of second language writing and spoken production, it has also been criticized against other more reliable measures such as a sentence-bases analysis (cf. Bardovi-Harlig, 1992).

³ For a thorough review of the impact of learning contexts and other aspects of learner development in L2 Spanish, see Lafford 2006.

(2004) investigated the gains obtained by 40 learners of Spanish in terms of oral fluency (measured by temporal and hesitation phenomena) and in oral proficiency based on the Oral Proficiency Interview (OPI) in two different contexts of learning: SA vs. AH. The results indicate that in some respects learners in the SA context made greater gains, both in terms of temporal and hesitation phenomena and in oral proficiency (OPI), than learners in the AH context.

The role of a domestic IM has been studied by Freed et al. (2004). They compared pretest and posttest measures of fluency from 28 learners of French in three learning contexts (i.e., IM, SA, and AH). The authors report that the domestic IM group made significant gains in oral performance with regard to the total number of spoken words, the length of the longest turn, the rate of speech, and speech fluency based on a composite of fluency measures. On the other hand, although participants in the AH and SA groups improved their overall speech fluidity, they achieved fewer gains than the IM group, which may be due to the fact that the IM learners were the ones that spent more time speaking and writing in French of all three groups analyzed in the study.

Although the previous studies have accounted for fluency gains in a highly detailed fashion, they present a methodological shortcoming in that they have not taken into account the range that may exist across different turns of speech. Specifically, all the studies cited in this literature review analyzed very short segments of learner production. For example, Temple (1992) analyzed four 4 minute-long segments (2 in the pre-test and 2 in the posttest) for each of the 22 learners of her corpus. Freed (1995) analyzed 45-second samples for each student's speech. Freed et al. (2004) increased the length of the segments to be analyzed by examining two 4 minute-long segments of each student's speech (two 1 minute-long segments from each student's pretest and posttest oral interviews). Given the breath of research done on L2 fluency, it is surprising that previous works have not advocated for the analysis of longer speech samples and multiple turns. The current study is designed to fill this gap in the literature by analyzing the 15 longest turns of oral production in learner speech for four groups of learners of L2 Spanish. Additionally, a group of 5 NSs of Spanish is included to provide a baseline comparison.

3. The Current Study

The following research questions guide the current study:

1. What ranges can be attested for a series of fluency measures for four groups of Spanish learners when larger samples of speech are taken into account?
2. What are the main differences in oral fluency among Spanish learners who have participated in a study abroad program (SA), an intensive overseas immersion (IM) and a formal academic context (AH)?
3. How do these ranges compare to those attested for NSs of Spanish?

3.1. Participants

A total of 25 participants volunteered to participate in the study (16 female, 9 male). All participants were recruited by the researcher through email. During the recording process, all participants signed the informed consent (approved by the Human Subjects IRB), completed the language and learning background questionnaires, and participated in a sociolinguistic interview with the researcher.⁴ There are two distinct groups: NSs of Spanish ($n=5$; M age = 28.6 years, $SD = 1.67$), and native English speaking learners of Spanish ($n=20$; M age = 19.7 years, $SD = 1.20$). Only one female reported in her language background questionnaire that in addition to English, she also considered Urdu and Panjabi as L1s. All 20 learners of Spanish reported that they had taken previous classes of Spanish at the University level (i.e., in the lower division) and that they were either already in the process of obtaining a major or a minor in Spanish or seriously considering this possibility. For the purpose of this study, the 20 learners were divided into 4 groups according to their Spanish learning experience:

⁴ A table with information on language experience is included in Appendix A.

AH₁ Group ($n=5$). This group consisted of 3 males and 2 females ranging in age from 18-19 years ($M = 18.8$, $SD = .446$); they were all enrolled in a Hispanic culture class (the second class that counts credit-wise towards the Spanish major or minor). All 5 participants averaged 4.8 years of pre-University classroom instruction and 2.6 semesters of University classroom instruction, and had never studied abroad.

AH₂ Group ($n=5$). This group consisted of 3 females and 2 males ranging in age from 19 to 21 years ($M = 20.4$, $SD = .89$). The 5 participants were enrolled in an introductory class in Hispanic Linguistics and had never studied Spanish abroad. All 5 participants had been enrolled in a number of University-level Spanish classes in the lower and upper divisions. They only studied Spanish in traditional at-home context, averaging 6 years of pre-university instruction and 5.6 semesters of University classroom instruction. These learners had never studied Spanish abroad.

IM Group ($n=5$). This is one of the two study abroad groups analyzed in this study. It consisted of 3 males and 2 females who participated in an intensive overseas immersion program in the north of Spain for 2 months in summer 2005. The 5 subjects ranged in age from 18 to 19 years ($M = 18.4$, $SD = .547$). They were all freshmen enrolled in a number of courses counting towards the Spanish minor or major (i.e., Hispanic culture, Hispanic literature introductory class and conversation). In addition to abiding by a no-English rule during their stay in Spain, these participants attended 4 hours of formal instruction each morning and approximately two hours of activities (i.e., sports, theatre or singing) each afternoon. These learners averaged 6 years of pre-University classroom instruction and 1 semester of University classroom instruction.

SA Group ($n=5$). This group consisted of 5 females ranging in age from 21 to 22 years ($M = 21.2$, $SD = .44$). All learners had participated in a study abroad program a Spanish-speaking country as part of their learning experience but with different duration. All these experiences included studies at the college level abroad. By the time the interview was recorded, they were taking an introductory class to Hispanic Linguistics (a more advanced class in the Spanish minor or major). These learners averaged 6 years of pre-University classroom instruction, 5.8 semesters of University classroom instruction, and 10.2 months studying Spanish abroad.

NS Group ($n=5$). This group consisted of 3 females and 2 males, all native speakers of Spanish, ranging in age from 26 to 30 years ($M = 28.6$, $SD = 1.67$). All participants were born in different regions of Spain: Seville, Cádiz, Valencia, and Basque Country. Although all members of this group indicated that their L1 was Spanish, one female reported in her language background questionnaire that she was raised bilingual in both Basque and Spanish. All were graduate students at a major research institution in the Midwest.

3.2. Data Collection

All participants completed a sociolinguistic interview with the researcher. A total of 54 questions guided the interview. They dealt with topics such as plans for vacations, a description of family members, if NSs or NNSs of a language make better language teachers, and how to prepare a special dish (cf. García-Amaya, 2009). Participants were interviewed individually. Sessions started with a presentation of the consent form and the language learning background questionnaire. All participants agreed to participate in the study. The interviews lasted between 30 and 40 minutes, participants were never stopped, and they answered every question for as long as they wanted. All transcriptions were digitally recorded, transcribed, and later analyzed according to 17 lexical and non-lexical elements as well as temporal variables.

3.3. Analysis

For the purpose of analysis, I selected the 15 longest turns from each participant's interview.⁵ I needed the longest samples of learner data in order to account for the range of fluency measures to be

⁵ In the present study, I adopt Levinson's (1983) definition of a turn, according to which "a turn is a time during which a single participant speaks, within a typical, orderly arrangement in which participants speak with minimal overlap and gap between them" (295-296).

considered in the current analysis (see below). This decision was based on one important criterion. Specifically, it was necessary to achieve a number of overall measures of fluency that included a higher range of learner oral production in their responses.

A table was prepared for each participant, and a total of 12 temporal and lexical measurements were calculated. Once the data had been processed, a mean was calculated for each measurement.⁶ This was done by dividing the sum of scores by the total number of turns ($n = 15$). The temporal and lexical measurements under analysis were the following:

1. *Total number of syllables per turn.* This measure (Temple, 1992) includes the number of syllables produced by each turn and includes syllables produced in Spanish and English words (i.e., *casa* – 2 syllables; *Steven* – 2 syllables). Filled pauses (i.e., *ah*, *ahm*, *eh*, *ehm*, *uh*, *uhm* & *mmm*) were not included in this measure.
2. *Total number of seconds per turn.* This measure (Temple, 1992) was obtained by measuring every turn in seconds. The software used was SoundScriber. This measure includes the time taken to produce the whole turn including words, filled pauses, and silent pauses.
3. *Rate of speech in syllables per second (inclusive).* Following previous investigations of speech rate (Binnenporte et al., 2005; Towell et al., 1996), the inclusive speech rate (i.e., total length of utterance including pauses) was calculated by dividing the total number of syllables produced in each turn (see 1, above) by the number of seconds taken to produce it. This measure also included all the syllables contained in repetitions, repairs and lexical fillers (i.e., discourse markers such as *bueno* “ok”, and *o sea* “in other words” among others).
4. *Rate of speech adjusted.* This measure was adapted from Freed et al. (2004), who measured speech rate in words per minute and excluded false starts, repetitions, partial repetitions, and items contained in repairs. In the current study, I measured syllables instead of words in order to account for the different lengths of words in Spanish (e.g., *verdaderamente*, truly, vs. *sí*, yes). The formula that was used is detailed here:

$$\text{rate of speech adjusted} = \frac{\text{syll.} - (\text{syll.in repeated segment} + \text{syll.in repaired segment})}{\text{seconds}}$$

5. *Filled Pauses.* Filled pauses (i.e., *uh*, *uhm*, *mmm*, *ah*, *ahm*, *eh* & *ehm*) were counted for every turn (Chafe 1980, Dechert 1980, Raupach, 1980, Freed et al. 2004)
6. *Repetitions.* This measure was derived from Lennon (1990) and Freed et al. (2004) and reflected all the repetitions that appeared in each turn. Repetitions were counted as units. For example: *cuando eh, cuando, estaba, cuando estaba a hacer mis, aplicaciones a la universidad ahm* “when, I was, when I was to do my, applications to the university ahm” would count as one repetition.
7. *Syllables in repetitions.* This measure reflected the syllables included in each repeated segment. For example, the repetition in *cuando eh, cuando, estaba, cuando estaba a hacer mis, aplicaciones a la universidad ahm* (repeated from 6, above) contained 7 syllables.
8. *Repairs.* This measure of fluency was derived from Lennon (1990) and Freed et al. (2004) and reflected the number of repairs as units used per turn. Examples of repair include grammar correction samples: *yo quiero ir la mes próximo, el próximo mes* “I want to go (the-fem) next month (masc. noun), (the-masc) next month (masc. noun) (= 1 self-correction).”
9. *Syllables in repairs.* This measure reflected the number of syllables included in each repaired segment, (i.e., *yo quiero ir la mes próximo, el próximo mes* would count as 5 syllables in the repaired segment).
10. *Total number of words.* This is a measure of general oral performance derived from Freed et al., (2004), who counted the total number of words spoken in a 2-minute interview sample. To achieve this measure, all the words produced in the 15 longest turns were summed.

⁶ To avoid error, the scores were calculated on raw scores divided by the total raw scores.

11. *Total words in English.* This measure is not normally reported in the literature but it is complementary to the previous measure and indicates words used in the L1. It includes names of towns: *Chicago, New York*, as well as lexical fillers (i.e., discourse makers): *like, you know, I mean*.
12. *Total number of syllables in English words.* This measure reflects the number of syllables included in English words produced in every turn.

4. Results

I begin this section with a summary of the results obtained through the analysis of the 25 transcriptions. Tables 1, 2, and 3 include the fluency measures for all participants in the five groups under analysis: AH₁, AH₂, IM, SA, and NS, as well as the range of measurements taken. The fluency measures of oral production which are reported as raw scores include: total spoken words, total syllables, total seconds, total filled pauses, total repetitions, total syllables in repetitions, total repairs, and total syllables in repairs. Table 1 also includes the ratio of filled pauses per syllables, the percentage of syllables in repetitions per total syllables, and the percentage of syllables in repairs per total syllables. Finally, both rates of speech, that is rate of speech inclusive and rate of speech adjusted, are also provided.

With regard to the total Spanish words, the range of total amount of spoken words increases from AH₁ (364-1585 words) to AH₂ (489-1609 words). A smaller range is attested for the IM group (556-1197 words), who had spent 2 months in Spain. The SA group, whose members had the longest experience in a Spanish-speaking country, is the group with the highest range of words produced in the segments analyzed (907-1619). In addition, the NSs show the highest range for the production of words (1831-2895).

As for the total number of syllables produced, there is a correspondence with the range for the total Spanish words. A gradual increase can be seen across the different groups of participants: AH₁ (529-2663), AH₂ (901-3102), IM (1035-2104), SA (1520-2876), and NS (3267-4958). As for total number of seconds per turn, there is an increase in the time learners take to talk about these 15 longest turns from AH₁ (247.9-1227) to the AH₂ (457.7-1144.6). Then, interestingly, the IM group shows the shortest range (447.4-539.5). The SA and the NS groups present very close ranges of seconds (557.7-1024.1 & 785-1029, respectively).

A comparison of both rate of speech inclusive and rate of speech adjusted yields very close measurements. The only difference between these two measurements was that the latter did not include syllables from repairs and repetitions. In fact, if a line were to be drawn joining all the results obtained for both rates of speech, it would most likely be very similar for both rates of speech, indicating that the amount of repetitions and repairs is not a robust strategy to gain fluency. The ratio of filled pauses per syllables becomes smaller as language experience (e.g., AH₁ & AH₂) and time abroad (IM & SA) increase, and while the ranges for the groups without experience abroad (i.e., AH₁ and AH₂) do not show much variation (1:5.69-1:22.13 and 1:7.41-1:25.43), the ratios of filled pauses per syllables are much smaller for the groups with experience abroad, IM and SA, (1:22.02-1:233.78 and 1:15.20-1:136.41, respectively), than for the AH groups (1:5.69-1:22.13 for AH₁ and 1:7.41-1:25.43 for AH₂). This being said, the NSs' range of filled pause per syllable (1:98.43-1:326.70) is closer to the IM's than to any other group of learners. With regard to percentage of syllables in repetitions per total syllables, it can be observed that again the NSs show rather distinctive values (0.20%-3.49%) than the learners. While the ranges for the AH₂ and the IM are close (1.77%-5.65% and 1.38%-4.01%, respectively), the AH₁ and the SA groups represent both ends of the continuum (1.40%-9.13% and 2.47%-6.25%, respectively). Again, the last measurement taken – that of percentage of syllables in repairs per total syllables – also yields very similar percentages for the groups with experience abroad along with the NS (0.23%-2.68% for IM, 0.60%-1.91% for SA, and 0.52%-2.03% for NS), separating both the AH groups to a different subgroup with ranges of 0.19%-4.66% for AH₁ and 1.41%-2.76% for AH₂.

Table 1. Fluency scores for the AH₁ and AH₂ groups

Group	Participant	Total Spoken		Total Seconds	Rate of Speech (Inc.)		Rate of Speech (Adj.)		Ratio Filled Pauses :		% Syll in Repetitions		% Syll in Repairs		
		Words	Syllables		Speech (Inc.)	Speech (Adj.)	Filled Pauses	Syllables	Repet	Total	per Total Syllables	Repair	Total	per Total Syllables	
AH ₁	John	727	1286	882.8	1.46	1.28	226	226	1 : 5.69	47	103	8.01	24	50	3.89
	Hannah	485	841	458.8	1.83	1.73	38	38	1 : 22.13	11	21	2.50	15	27	3.21
	Helen	364	643	413.9	1.55	1.49	83	83	1 : 7.75	8	9	1.40	7	19	2.95
	Bill	1585	2663	1227.6	2.17	1.87	301	301	1 : 8.85	91	243	9.13	37	124	4.66
	Peter	330	529	247.9	2.13	1.96	36	36	1 : 14.69	18	41	7.75	1	1	0.19
Range for AH ₁		364-1585	529-2663	247.9-1227	1.46-2.17	1.28-1.96	38-301	38-301	1 : 5.69 - 1 : 22.13	8-91	9-243	1.40-9.13	1-37	1-124	0.19-4.66
AH ₂	Daniel	501	912	621.3	1.47	1.36	123	123	1 : 7.41	30	43	4.71	11	23	2.52
	Sara	489	901	457.7	1.97	1.88	91	91	1 : 9.90	12	17	1.89	10	22	2.44
	Ryan	1345	2409	915.7	2.63	2.45	199	199	1 : 12.11	85	136	5.65	16	34	1.41
	Rebecca	1095	1813	638.1	2.84	2.71	37	37	1 : 4.9	17	32	1.77	17	50	2.76
	Eve	1609	3102	1144.6	2.71	2.59	122	122	1 : 25.43	40	88	2.84	21	45	1.45
Range for AH ₂		489-1609	901-3102	457.7-1144.6	1.47-2.84	1.36-2.71	37-199	37-199	1 : 7.41 - 1 : 25.43	12-85	17-136	1.77-5.65	10-21	22-50	1.41-2.76

Table 2. Fluency scores for the IM and SA groups

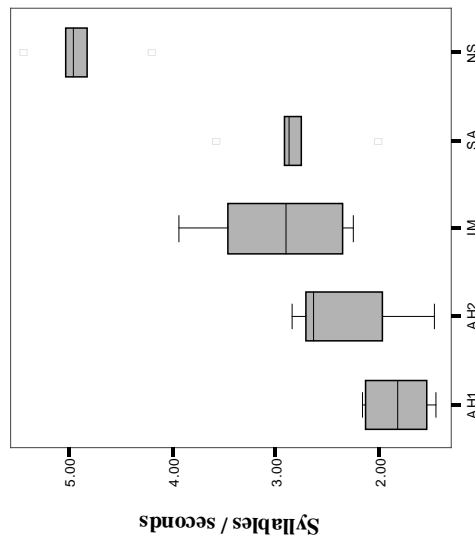
Group	Participant	Total Spoken		Total Seconds	Rate of Speech (Inc.)		Rate of Speech (Adj.)		Ratio Filled Pauses :		% Syll in Repetitions		% Syll in Repairs		
		Words	Syllables		Speech (Inc.)	Speech (Adj.)	Filled Pauses	Syllables	Repet	Total	per Total Syllables	Repair	Total	per Total Syllables	
IM	David	1197	2104	534.5	3.94	3.75	9	9	1 : 233.78	39	61	2.90	13	37	1.76
	Patrick	704	1302	447.4	2.91	2.86	6	6	1 : 217.00	6	18	1.38	1	3	0.23
	Kim	704	1245	528.3	2.36	2.23	47	47	1 : 26.49	17	40	3.21	11	26	2.09
	Matt	556	1035	459.9	2.25	2.15	47	47	1 : 22.02	9	22	2.13	10	25	2.42
	Kathryn	1062	1868	539.5	3.46	3.23	14	14	1 : 133.43	32	75	4.01	14	50	2.68
Range for IM		556-1197	1035-2104	447.4-539.5	2.25-3.94	2.15-3.75	6-47	6-47	1 : 22.02 - 1 : 233.78	6-39	18-75	1.38-4.01	1-14	3-50	0.23-2.68
SA	Mila	1612	2830	1024.1	2.76	2.57	103	103	1 : 27.48	97	177	6.25	4	17	0.60
	Carrie	1619	2876	1002.3	2.87	2.75	142	142	1 : 20.25	46	71	2.47	23	47	1.63
	Heather	1449	2319	656	3.54	3.4	17	17	1 : 136.41	40	71	3.06	6	15	0.65
	Maryl	939	1630	557.7	2.92	2.76	28	28	1 : 58.21	22	44	2.70	18	49	3.01
	Liz	907	1520	771.5	1.97	1.88	100	100	1 : 15.20	27	41	2.70	8	29	1.91
Range for SA		907-1619	1520-2876	557.7-1024.1	1.97-3.54	1.88-3.4	17-142	17-142	1 : 15.20 - 1 : 136.41	22-97	41-177	2.47-6.25	4-23	15-49	0.60-1.91

Table 3. Fluency scores for the NS group

Group	Participant	Total Spoken Words	Total Syllables	Total Seconds	Rate of Speech (Inc.)	Rate of Speech (Adj.)	Total Filled Pauses	Ratio Filled Pauses : Syllables	Total Repet	Total Syll Repet	% Syll in Repetitions	Total Repair	% Syll in Repairs
NS	Manuel	2353	4134	822	5.03	4.8	42	1 : 98.43	69	106	2.56	35	2.03
	Victor	1831	3267	785	4.16	3.96	10	1 : 326.70	66	114	3.49	19	1.44
	Irene	2537	4486	830	5.4	5.22	27	1 : 166.15	26	61	1.36	31	2.01
	Elsa	2504	4415	891	4.96	4.82	28	1 : 157.68	47	92	2.08	9	0.61
	Luisa	2895	4985	1029	4.82	4.78	7	1 : 712.14	8	10	0.20	8	0.52
Range for NS		1831-2895	3267-4988	785-1029	4.16-5.40	3.96-5.22	7-42	1 : 98.43 - 1 : 326.70	8-69	10-114	0.20-3.49	8-26	0.52-2.03

Figure 1 : Rate of speech (inclusive) for all participant groups

Rate of speech for all participant groups



Learners grouped according to previous learning experiences

Figure 2: Linear regression for rate of speech (inclusive) by total spoken words produced in the 15 longest turns

Syllables per second by total spoken words

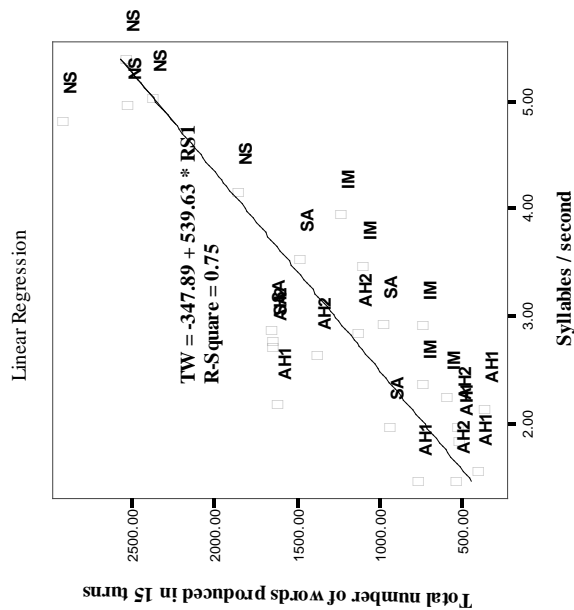


Figure 1 presents a box plot that graphically depicts ranges of rate of speech (inclusive) for all participants groups and clearly shows how the groups of learners with a higher rate of speech compare to those with experience abroad (i.e., IM & SA). It is also noteworthy that the median for both IM and SA groups is the same although the ranges indicate that the rate of speech of IM learners is somewhat higher than that of SA learners (2.25-3.94 vs. 1.97-3.54 syllables per second, respectively). This figure also shows how the participants in the AH₂, who at the time when the data of the present study were collected were taking the same class as the participants in the SA group, have a slower range of rate of speech (1.47-2.84 syllables per second). The NS group and the AH₁ group show both ends of the distribution, the former showing a range that is not reached by any learner in the study, and the latter showing the smallest range of all groups.

Figure 2 presents a scatter plot with a linear regression for the rate of speech inclusive by total spoken of words produced in the 15 longest turns for all 25 participants in the study.⁷ The scatter plot shows that the NSs are distributed very closely. It also shows that the NSs have higher rates of speech and higher production of words in the longest 15 turns than the learners. Regarding the learner groups, the AH₁ group shows a close distribution, with the exception of one speaker, who has a similar rate of speech to his cohorts but a word production closer to that of the AH₂ and SA speakers. The AH₂ and IM groups show more variation than the AH₁ and the NS groups for rate of speech, but are more uniform in their production of words. Overall, it can be stated that there is considerable overlap among the four learner groups. In fact, there are learners from different groups with almost identical values.

Regarding the statistical analysis, a one-way ANOVA and Tukey-HSD post-hoc test were performed for the 5 speaker groups as independent variables. Rate of speech inclusive, rate of speech adjusted, total number of words produced, and total number of filled pauses produced were the dependent variables.⁸ One-way ANOVA was chosen because there was one observation for each of the independent variables. This statistical procedure was only performed for those fluency measures for which raw scores data could be provided. Each ANOVA test yielded a significant result: rate of speech ($F(4,20) = 22.622, p < .000$; effect size = .81); rate of speech adjusted ($F(4,20) = 24.465, p < .000$; effect size = .83); number of words ($F(4,20) = 13.387, p < .000$; effect size = .72); and the production of filled pauses ($F(4,20) = 3.361, p < .029$; effect size = .40).

With regard to rate of speech, the post-hoc tests (multiple comparisons) indicated a significant difference for rate of speech between the AH₁ and the IM ($p = .023$), as well as for the NSs and the remaining learner groups ($p = .000$). With regard to rate of speech adjusted, the AH₁ and IM were shown to be significantly different from one another ($p = .015$), as well as the AH₁ and the SA ($p = .045$). Once again, the test showed that the NSs were significantly different from all learner groups ($p = .000$). With regard to the total number of words, the NSs were also significantly different from all learner groups: AH₁ ($p = .000$), AH₂ ($p = .000$), IM ($p = .000$), and SA group ($p = .004$). Lastly, the post-hoc tests showed that the production of filled pauses did not have a significant difference across groups.

5. Discussion

The main goal of the study was to show the ranges for various dimensions of fluency in three different learning environments (i.e., AH, IM, and SA) by 4 groups of learners. Of the 4 different groups of learners, two consisted of learners without experience abroad (i.e., AH₁ and AH₂), while the other two (i.e., IM and SA) had experience abroad. Finally, a group of NSs was also included in the design to serve as a baseline for all the temporal measures analyzed in the study. It is important to recall that the following measurements were taken for each of the 15 longest turns of speech produced by each participant: total number of spoken words, total number of syllables, rate of speech inclusive and adjusted (without repetitions and repairs), ratio of filled pauses per syllables, percentage of syllables in repetitions per total number of syllables, and percentage of syllables in repairs per total number of syllables.

⁷ The R^2 value for the regression is .075.

⁸ An ANOVA was preferred to MANOVA (multivariate analysis of variance) because the sample size was too small ($n = 25$). Tukey-HSD was used to adjust the p -values to control for multiple comparisons in the post-hoc tests. Tukey-HSD provides a more accurate correction of the p -value (to maintain an overall $\alpha = .05$) than Bonferroni, which is overly conservative.

In light of the research questions, the most important finding of this study is the range exhibited for each of the fluency measures that was calculated for each participant group. As was shown in Tables 1, 2, & 3 for a number of speakers, the range of scores for members of the same group show that in terms of fluency measures, some group members do not conform to the pattern established for the rest of the group. Previous fluency studies that analyzed very small speech fragments as representative for all the speech of individual learners may have missed important generalizations for the learner groups under study. Based on the results from the current study, it would seem that in order to account for fluency from a broader perspective, it is necessary to address the existing variation in temporal variables in different moments of the conversation in a systematic way. Analyzing larger samples of speech seems to be a very appropriate way of approaching this need.

With regard to the second focus of the study, that is, fluency measures for different groups of learners, the measures of fluency obtained in the analysis indicate that the relationship between different measures of fluency across different levels of instruction in University-level Spanish (i.e., AH₁, AH₂ & SA) varies depending on the measure and the level of proficiency.⁹ The results also indicate that the SA group surpasses the AH₁ group in all the aforementioned fluency measures. Nevertheless, comparison between the AH₂ and the SA group yields unexpected results, mainly because all 10 learners in both groups were taking the same class at the time when the interviews were recorded. In general, there is little question that the SA group members are able to produce more words (907-1619 vs. 489-1609) and speak at a faster rate (1.97-3.54 vs. 1.47-2.84). However, distances among groups become shorter in terms of filled pauses (17-142 vs. 37-199), repetitions (22-97 vs. 12-85) and repairs (4-23 vs. 10-20), where the SA group shows little differences in comparison to the AH₂.

These results are consistent with previous accounts reported in the literature (i.e., Freed et al., 2004; Segalowitz and Freed, 2004), where the learners with experience abroad were the more “fluent” (i.e., faster rate of speech, larger amount of words, fewer filled pauses and fewer repetitions and repairs). With regard to the study abroad experience in general, both the IM and SA groups had two rather different experiences which reflect their different outcomes.¹⁰ Although the IM group only spent 2 months in an overseas Spanish speaking environment (as opposed to the SA learners who spent an average of 10.2 months abroad), the results show that while the SA learners can produce longer turns (more words and more time holding the floor), the IM learners have a higher rate of speech, fewer repetitions, fewer repairs and fewer filled pauses. It is worth recalling that Freed et al. (2004) concluded that out of the three groups that they had included in their study (i.e., AH, IM and SA), the domestic IM context was the most successful in terms of rate of delivery. However, it should be recalled that the overseas IM learners in the current study spent their immersion experience in Spain, whereas the students included in the study by Freed et al.’s study were involved in a domestic immersion program in the United States. Another difference between the two IM groups is that those under analysis in the current study were immersed in a foreign Spanish-speaking environment for 2 months, while the domestic IM learners were immersed in a second language environment in a year-round program.

In terms of the ANOVA and post-hoc tests that were run, it was shown that all learner groups behaved statistically different from the NS group. Nevertheless, certain important differences were found among learner groups. In terms of rate of speech, only the AH₁ group was significantly different from the IM group for rate of speech ($p = .023$). Also, differences between the AH₁ group and the IM and SA groups were found to be significant for rate of speech adjusted ($p = .015$ and $p = .045$, respectively). The AH₂ group, however, was not found to be statistically different from the two advanced groups for these measures. Although this may seem to imply that no differences exist between AH₂ group and the groups of learners that had been abroad, it should be recalled that the

⁹ The IM group learners are not considered to discuss this first research question because they entered college very recently and did not have the opportunity to take a variety of courses in Spanish. The other learners, on the other hand, had the opportunity to enroll in other classes (i.e., in the lower division) before pursuing their major or minor (i.e., in the upper division).

¹⁰ Neither the IM nor SA groups were tested before their experience abroad. Thus, the comparisons presented on this study correspond to their oral performance after being abroad.

range of scores for rate of speech for the AH₂ group, when compared to the IM and SA group, was considerably lower.

6. Conclusions and Future Directions

The empirical results of this investigation have made further contribution to the study of second language fluency. From a methodological standpoint, it has demonstrated that measuring different turns is a more accurate way of calculating overall fluency measures for second language learners. Given the amount of intrarange variation appearing within each group of learners, the methodology advocated in the current study seems to be the most insightful for the research goals. With regard to the research questions that drove this study, this paper first sought to investigate the ranges for a series of oral fluency measures for four groups of L2 Spanish learners when larger samples of speech were taken into account. The results showed that between-learner variation on a number of fluency measures makes it difficult to distinguish between groups with experience abroad, and groups without experience abroad.

In this sense, little differences were found for spoken words, total number of seconds per turn, percentage of syllables in repetition per total number of syllables, and percentage of syllables in repair per total syllables between the AH₂, and the SA groups. On the other hand, with regard to both inclusive and adjusted rate of speech, it was found that syllables from repetitions and repairs did not make a large difference for these learner groups. The second research question sought to investigate the differences in oral fluency among Spanish learners who had participated in a study SA, an intensive overseas IM, and a formal AH academic context. The results indicated that no learner group proved to be superior with regard to all fluency measures. Much variability was observed in this respect, and while the IM group was superior to the other three groups on rate of speech (inclusive and adjusted), ratio of filled pauses per syllables, and percentage of syllables in repetition per total syllables, the SA group was superior than the IM group on number of spoken words, total number of seconds per turn, and percentage of syllables in repair per total syllables. Also on number of spoken words and total number of seconds per turn, little difference was found between the AH₂ and SA groups.

The last research question addressed the ranges between the groups of L2 Spanish learners with NSs of Spanish. The NS group surpasses the learner groups and thus is not comparable to them for: spoken words, rate of speech (both inclusive and adjusted), and the ratio of filled pauses per syllables. Moreover, the one-way ANOVA showed that the NSs were significantly different from all learner groups in both rates of speech (inclusive and adjusted), as well as total number of words. They IM group is the only learner group that seems to be comparable to the NSs on the percentage of syllables in repetition per total syllables as well as on the percentage of syllables in repair per total syllables. As it can be seen, these findings break down the superficial generalization of having a natural break in fluency between the AH groups, the SA/IM groups and the NS group.

With regard to future research, an investigation into the differences between different types of immersion programs (that is, domestic vs. overseas) could shed more light into the effects of fluency of this particular type of learning environment. Although the specialized literature has addressed the role of immersion programs, the role of overseas immersion programs such as the one described in the current investigation has not been studied extensively. Indeed, one aspect of IM program mentioned in the current study is that it takes place in an abroad setting. However, it is not exactly like a study abroad program (as traditionally described in the literature), since its students are required to abide by specific rules prohibiting the use of English while abroad. In this investigation, 3 of the participants in the SA group reported that although they spoke mostly Spanish during their experience abroad, they also used English to talk to their American friends, a practice that hardly ever takes place in the IM program described here, where the learners sign a document declaring their pledge to the program rules, and in particular, to the no-English rule. Additionally, even more significant is the fact that the intensive overseas immersion program behaves in a similar way to the intensive domestic immersion program reported by Freed et al. (2004). This opens a new line of research to compare these two different immersion learning contexts.

Finally, there is little doubt that, given the findings of the current study, the future of fluency research, in terms of its impact on SLA, is a promising field of research. More research is necessary in

order to understand fluency strategies and what elements allow learners to develop in their oral productions. The current investigation has shown that more rigorous methodologies are necessary in order to gain access into subtle differences across learners. In doing so, future work should continue in the line of data analysis advocated here, since it allows for a more precise description of learner data and interlanguage development in general.

Acknowledgments

I am indebted to Kathleen Bardovi-Harlig for the invaluable insight and guidance that she offered throughout the research process. I am also grateful to Kimberly Geeslin and two anonymous reviewers for their help with previous versions of this paper. I am thankful to Stephanie Dickinson for her help with the statistical analysis. Of course, all shortcomings are my own.

Appendix: Participants' Language Experience

Table 4. Language learning experience of Spanish as an L2

Group	Participant	Years of pre-University classroom instruction	Semesters of University classroom instruction	Months spent abroad
AH ₁	John	6	3	0
	Hannah	6	3	0
	Helen	4	1	0
	Bill	4	3	0
	Peter	4	3	0
	Average	4.8	2.6	0
AH ₂	Daniel	6	7	0
	Sara	6	4	0
	Ryan	6	7	0
	Rebecca	6	6	0
	Eve	6	4	0
	Average	6	5.6	0
IM	David	6	1	2
	Patrick	6	1	2
	Kim	5	1	2
	Matt	7	1	2
	Kathryn	6	1	2
	Average	6	1	2
SA	Mila	6	4	12
	Carrie	6	7	13
	Heather	6	7	12
	Maryl	6	4	7
	Liz	6	7	7
	Average	6	5.8	10.2

References

- Bardovi-Harlig, K. (1992). A second look at T-unit analysis. *TESOL Quarterly*, 26, 390-395.
- Binnenporte, D., Van Bael, C., den Os, E., Boves, L. (2005). Gender in everyday speech and language: A corpus-based study. *Interspeech*, 1-4.
- Chafe, W. L. (1980). Some reasons for hesitating. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 169-180). The Hague: Mouton.
- Dechert, H. W. (1980). Pauses and intonation of verbal planning in second-language speech productions: Two examples from a case study. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 271-285). The Hague: Mouton.
- Dechert, H. W. & Raupach, M. (1987). *Psycholinguistic models of production*. Norwood, NJ: Ablex.
- DeKeyser, R. (1991). Foreign language development during a semester abroad. In B. F. Freed (Ed.), *Foreign language acquisition and the classroom* (pp. 104-118). Lexington, MA: D.C. Heath.
- Freed, B. F. (1990). Language learning in a study abroad context: the effects of interactive and non-interactive out-of-class contact on grammatical achievement and oral proficiency. In J. Atlatis (Ed.), *Linguistics, language teaching and language acquisition: the interdependence of theory, practice and research (GURT 1990)* (pp. 459-477). Washington, DC.: Georgetown University Press.
- Freed, B. F. (1993). Assessing the linguistic impact of study abroad: what we currently know-what we need to learn. *Journal of Asian Pacific Communication*, 4, 151-66.
- Freed, B. F. (1995). What makes us think that students who study abroad become fluent? In B. F. Freed (Ed.), *Second language acquisition in a study abroad context*, (pp. 123-148). Philadelphia: John Benjamins.
- Freed, B. F. (1998). An overview of issues and research in language learning in a study abroad setting. *Frontiers*, 4, 31-60.
- Freed, B. F., Segalowitz, N., & Dewey D. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26, 275-301.
- García-Amaya, L. (2008). The effect of topic on rate of speech. *ELIA*, 8, 117-150.
- Goldman-Eisler, F. (1964). Discussion and further comments. In E. H. Lenneberg (Ed.), *New directions in the study of language* (pp. 109-131). Cambridge, MA: MIT Press.
- Hunt, K.W. (1965). *Grammatical structures written at three grade level* (Research Report No. 3). Urbana, IL: National Council of Teachers of English.
- Lafford, B. A. (2006). The effects of study abroad vs. classroom contexts on Spanish SLA: Old assumptions, new insights an future research directions. In Klee, C.A., & Face, T. L. (Eds.), *Selected proceedings of the 7th conference on the acquisition of Spanish and Portuguese as first and second languages* (pp. 1-25). Somerville, MA: Cascadilla Proceedings Project.
- Levinson, S. (1983). *Pragmatics*. Cambridge, England: Cambridge University.
- Olynyk, M., d'Anglejan, A., & Sankoff, D. (1990). A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. In R. Scarcella, R. Andersen & S. Krashen (Eds.), *Developing communicative competence in a second language* (pp. 139-155). Rowley, MA: Newbury House.
- Raupach, M. (1980). Temporal variables in first and second language speech production. In H. W. Dechert & M. Raupach (Eds.) *Temporal variables in speech* (pp. 49-60). The Hague: Mouton.
- Sajavaara, K. (1987). Second language speech production: Factors affecting fluency. In H. W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 45-65). Norwood, NJ: Ablex.
- Segalowitz, N. & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26, 173-199.
- Segalowitz, N., Freed, B. F., Collentine, J. G., Lafford, B. A., Lazar, N., & Díaz-Campos, M. A. (2004). A comparison of Spanish second language acquisition in two different learning contexts: study abroad and the domestic classroom. *Frontiers* 10, 21-38.
- Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14, 357-385.
- Temple, L. (1992). Disfluencies in learner speech. *Australian Review of Applied Linguistics*, 15, 29-44.
- Towell, R. Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, 17, 84-11.

Selected Proceedings of the 11th Hispanic Linguistics Symposium

edited by Joseph Collentine,
Maryellen García, Barbara Lafford,
and Francisco Marcos Marín

Cascadilla Proceedings Project Somerville, MA 2009

Copyright information

Selected Proceedings of the 11th Hispanic Linguistics Symposium
© 2009 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-432-4 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

García-Amaya, Lorenzo. 2009. New Findings on Fluency Measures across Three Different Learning Contexts. In *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, ed. Joseph Collentine et al., 68-80. Somerville, MA: Cascadilla Proceedings Project.

or:

García-Amaya, Lorenzo. 2009. New Findings on Fluency Measures across Three Different Learning Contexts. In *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, ed. Joseph Collentine et al., 68-80. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2203.