TESL-EJ Teaching English as a Second or Foreign Language
The Electronic Journal for English as a Second Language

March 2011—Volume 14, Number 4

Improving Speaking Fluency for International Teaching Assistants by Increasing Input

Greta J. Gorsuch
Texas Tech University
<greta.gorsuch@ttu.edu>

Abstract

One challenge for many international teaching assistants (ITAs) is improving their spoken English fluency after arrival in the U.S.A. It may be argued that poor fluency, with its hallmarks of slow speech rate, false starts, and particularly pauses that violate phrasal boundaries, account for the failure of many ITAs to be certified by their institutions to teach undergraduate labs or classes. And, for a variety of reasons, simply being in graduate school in the U.S.A. may not result in ITAs rapidly improving their English, even after a semester or more of specialized ITA courses. This study explores the question of whether an input approach, in addition to a production-oriented ITA preparation approach, will improve ITAs' spoken fluency in post-treatment teaching simulations. In this study, 28 participants in an ITA preparation course engaged in twice-a-week repeated reading (RR) sessions, in which they repeatedly and silently read 500-word basic popular science texts along with an audio recorded model of the text. In this study, ITAs' gains in reading fluency and comprehension were tracked throughout an academic semester for a total of twenty 30-minute RR sessions. In addition, two ITAs' teaching simulation presentations were audio recorded, once at the beginning and once at the end of the fourteen-week course. Changes in speech rate, percentage of grammatically intact pause groups, and percentage of disfluent pause groups were tracked. ITAs' reading fluency and comprehension increased significantly, while the percentage of intact pause groups increased, and the percentage of "split" pause groups decreased. While causality between the RR treatments and improvements in ITAs' spoken fluency cannot be strictly stated, a theoretical model of how extensive input may promote speaking fluency is presented, along with specific suggestions on creating input-focused programs for ITAs.

Introduction

One challenge for many international teaching assistants (ITAs) is improving their spoken English fluency after their arrival in the U.S.A. (Gorsuch, 2008; Heidish, 2006). It may be argued that poor fluency, with its hallmarks of slow speech rate, false starts, and particularly

pauses which violate phrasal or clausal boundaries, account for the failure of many ITAs to be certified by their institutions to teach undergraduate labs or classes. And, for a variety of reasons, simply being in graduate school in the U.S.A. may not result in ITAs rapidly improving their English, even after a semester or more of specialized ITA courses.

The current study explores the question of whether an input approach, in addition to a more conventional production-oriented ITA preparation approach, will improve ITAs' spoken fluency in post-treatment teaching simulations. In other words, if ITAs notice the features of fluent English speech, such as grammatically intact pause groups in input that is comprehended, in addition to being taught about them (as arguably may done in conventional ITA courses), it might be hypothesized that ITAs' improvement on these features will be reflected in their continuous, non-conversational speech (such as a teaching simulation). In this study, 28 participants in an ITA preparation course engaged in twice-a-week repeated reading (RR) sessions, in which they repeatedly and silently read 500-word basic popular science texts along with an audio recorded model of the text.

Literature Review

Speaking fluency and International Teaching Assistants. L2 speaking fluency is an area of persistent interest to applied linguists (Kopenen & Riggenbach, 2000). There are a number of key areas of interest, such as what hesitation and self repair phenomena tell us about L2 learners' processing in syntactic and discourse contexts (e.g., Deese, 1980; Ejzenberg, 2000; Lennon, 2000; Pawley & Syder, 2000; Riggenbach, 1989); how pauses, and clause and phrase boundaries are related in fluent and non-fluent L1 and L2 speakers (e.g., Butterworth, 1980; Crystal & Davy, 1969; Goldman Eisler, 1968; Levelt, 1989; Pawley & Syder, 2000) and how these may be measured to show L2 learner fluency development over time (Lennon, 2000); and how L2 speakers' pauses affect listeners' perceptions of an L2 speaker's fluency and ability to communicate ideas (e.g., Butcher, 1980; Ejzenberg, 2000; Olynak, Anglejan, & Sankoff, 1990; Wennerstrom, 2000). Whether one defines fluency in a broad sense ("semantic density, sociolinguistic appropriateness, and creativity in language use," Kopenen & Riggenbach, 2000, p. 7) or a narrow sense ("the speed and smoothness of oral delivery, "Lennon, 2000, p. 25) there is no doubt that speaking fluency is implicated in judgments of whether an L2 speaker has communicative competence (Olynak, Anglejan, & Sankoff, 1990; Pawley & Syder, 1983; Riggenbach, 1989). This study focuses on fluency defined in the "narrow" sense.

ITA educators have been concerned with the speaking fluency of the Chinese, Korean, and Indian biology, chemistry, physics, and engineering graduate students who have comprised their main clientele for the past thirty years (Ford, Gappa, Wenddorff, & Wright, 1991; Gorsuch, 2003; Pialorsi, 1984; Smith, Byrd, Nelson, Barrett, & Constantinides, 1992). These upper-intermediate to advanced L2 English speakers must use their L2 to teach undergraduate classes in their disciplines to be supported as graduate students. Depending on the English language learning backgrounds of ITAs, the task of teaching content and managing undergraduate classrooms places extraordinary demands on ITAs' L2 communicative competence (Gorsuch, 2008).

According to commentary on speaking fluency, L1 and L2 speakers' ability to manipulate lexis, syntax, and larger discourse structures, and access "fluent chunks" (memorized

formulaic phrases and sentences) without false starts, undue hesitations, and violations of clause boundaries depends greatly on speaking task complexity and familiarity (Pawley & Syder, 2000; Riggenbach, 1989). The task demands of teaching in a second language, particularly for ITAs without teaching experience even in their L1s, guarantee formidable challenges to ITAs' speaking fluency in English (for a description of skills needed to give a monologic speech, see Dudley-Evans & St. John, 1998). Because of this, and because of ITA educators' state- and institution-mandated sensitivity to the needs of ITAs' audiences (U.S. undergraduates) (Hoekje & Linnell, 1994; Thomas & Monoson, 1991) fluency receives attention in the assessments and instruction used in ITA education programs. Early ITA performance assessments included scoring criteria on "delivery" (e.g., Hinofotis, Bailey, & Stern, 1981) and current assessments include scoring criteria on "speech flow" and fluency (Gorsuch, 2006a; Gorsuch, Meyers, Pickering, & Griffee, 2010; Smith, Meyers, & Burkhalter, 1992). Further, ITA educators suggest their ITA preparation courses include fluency (e.g., Heidish, 2006; Papajohn, Alsberg, Bair, & Willenborg, 2002), and publish materials which focus partly on spoken fluency (Byrd, Constantinides, & Pennington, 1989; Gorsuch et al., 2010; Smith et al., 1992).

How fluency is operationalized in this study. There are five features of fluency focused on in this study: Grammatically intact pause groups ("fluent" pause groups); pause groups which violate phrase or clause boundaries ("split" pause groups); pause groups with self-repairs involving two or more utterances ("false start" pause groups); one or two word fillers without semantic meaning ("fillers"); and rate of speech.

In spoken English, pauses occur most of the time at phrasal or clausal boundaries, and at the ends of sentences (Butcher, 1980; Crystal & Davy, 1969; Goldman Eisler, 1968; Pawley & Syder, 1983, 2000). Pauses are defined as periods of silence of .2 seconds or longer (Butterworth, 1980). Pauses that violate phrase or clause boundaries do occur in the extemporaneous monologic talk of native speakers of English, but when speakers engage in nonverbatim rehearsal several times, increasing amounts of pauses occur at phrase or clause boundaries (Butterworth, 1980). Thus pauses are external expressions of more and less automaticity in cognitive planning and processing of speakers (Butterworth, 1980; Pawley & Syder, 2000). Examples of fluent pause groups from this study are *uh so before I can get to this // I would like to talk briefly about different types of bio fuels // that can be used // as energy sources*. Examples of split pause groups are *uh this is more commonly known as // digestion* and *I have been interested in // alternative energy for quite a while*. In this study, the term "pause groups" is taken to mean the same as "thought groups" mentioned in the Materials section.

False start pause groups and fillers also suggest the flows and ebbs of a speaker's cognitive processing in the face of on-line language production demands. False starts or self-repairs embedded in or comprising pause groups may suggest a struggle involving "the speaker's concern for difficulties arising in the use of language" and their simultaneous "attempts to maintain contact with their interlocuters" (Olynack et al., 1990, p. 141). Both fillers and false starts may be a compensatory strategy to maintain a fluent delivery (Oppenheim, 2000) by buying processing time and slowing down production demands (Crystal & Davy, 1969). Examples of false start pause groups taken from this study are *uh if there's are there any*

questions and *so what the DNA does it DNA it*. Examples of fillers are *um* and *mmm* and *uh ok*.

Rate of speech is of interest to SLA researchers such as Ellis (1990) who are concerned that L2 learners who focus on increasing fluency may do so at the expense of accuracy. Ellis conceives of fluency as a learners' control over "channel" and thus operationalizes fluency as speech rate ("the number of syllables produced in one minute of speech") (1990, p. 87). L1 speakers have higher rates of speech than L2 speakers due to higher automaticity with speech production processes (Ejzenberg, 2000; Wiese, 1984). Rate of speech is operationalized in this study as words per minute after fillers, repetitions, self corrections, and parts of words are discounted.

ITA speaking development as a second language acquisition issue. The current study proposes that ITAs' speaking development, including improvement in temporal aspects of fluency (see discussion above), be viewed through the lens of second language acquisition (SLA). The SLA "problems" of ITA candidates at U.S. university campuses are multi-faceted. As early-career graduate students, many ITAs do not have time to devote to English language study. They may feel that courses within their disciplines are more important than ESL courses (Gorsuch, 2006b). Even if ITAs attend specialized ESL courses, it will not have enough impact if the courses meet only once or twice a week (Griffee, Gorsuch, Britton & Clardy, 2009). Further, ESL classrooms may not be particularly acquisition-rich (see Ellis, 1997 for a penetrating discussion). ITAs may also limit themselves to interacting with colleagues who share their L1s and thus have little contact with English speakers (Gorsuch, 2008; Petro, 2006). This reticence is thought to negatively affect language acquisition (Asker, 1998).

An earlier study on discipline-specific practica (Gorsuch, 2006b) argued for ITAs to have extensive, guided opportunities to observe and participate in U.S. undergraduate labs and classes. Without such programs ITAs will not have access to context-specific L2 comprehensible input that is taken to be a prerequisite for L2 acquisition. However, even in such "successful" practica (where ITAs participated in at least one undergraduate class or lab per week for an entire semester), ITAs varied in their success in attending to input (Gorsuch, 2006b).

Nonetheless, ITAs' needs for L2 acquisition remain the same as for any other L2 learner group. ITAs need comprehensible input (Krashen, 1985), and they need to attend to that input (Ard, 1987; Schmidt, 1993). Given that spoken fluency in an L2 "may take the adult learner of a foreign language years to achieve" (Pawley & Syder, 1983, p. 199), and given that ITAs must use reasonably fluent speech to teach in a setting where the L2 is arguably hard to acquire, ITAs and ITA educators need to formulate and test remedies for developing this important component of communicative competence. In two early fluency scholars' words, communicative competence must take into account a description of fluency, which includes the ability to regularly encode "whole clauses, in their full lexical detail, in a single encoding operation" thus avoiding "mid-clause hesitations" (Pawley & Syder, 1983, p. 204).

Conventional responses to ITA speaking development. Many descriptions of ITA programs treat the development of ITAs' speaking abilities as a largely output-focused process. There are depictions of courses which improve segmental and suprasegmental aspects of

pronunciation (Alsberg, 2002; Chesser, Carroll, Macero, & Tice, 2002; Cornell University, 2007; Gorsuch, Stevens, & Brouillette, 2003; Panvini, 2002; Papajohn et al., 2002; Zukowski-Faust, 1984), and also facility with communicative functions (Smith, Byrd, Constantinides, & Barrett, 1991). Some articles include developing ITAs' metacognition for speaking (Alsberg, 2002; Papajohn et al., 2002). Other descriptions include instructor and U.S. undergraduate feedback on ITA production in microteaching sessions (Ford et al., 1991; International TA Program, 2009; Papajohn et al., 2002; Tanner, 1991). ITA education commentators and textbook writers echo this production-focused orientation (Byrd & Constantinides, 1995; Byrd et al., 1989; Dickerson, 2002; Madden & Myers, 1994; Gorsuch et al., 2010; Smith et al., 1992).

On the face of it, such production-oriented instruction may find its inspiration in skills theories in SLA, which posits fluency in terms of accessing and using speech production mechanisms (Butterworth, 1980; Deese, 1980; Lennon, 2000; Segalowitz, 2000) which learners have more or less automaticity with. Certainly instructional features within a speaking production orientation, such as explicit instruction and controlled practice in suprasegmental aspects of pronunciation, speech rehearsal, and feedback on spoken performance with an aim to enhance metacognitive awareness of and control over performance, might be supported by SLA skills theories. But this may be an oversimplification, and a production orientation may simply be robust pedagogical beliefs about having ITAs directly engage in the areas of most visible weakness. However, as noted above, ITA courses may not meet often enough, or for long enough, to bring about realistic changes in speech production skills.

In contrast, some literature of ESP (English for Specific Purposes), from which ITA education arguably draws inspiration, emphasizes authentic materials for learner L2 input, as well as emphasizing L2 production (e.g., Dudley-Evans et al., 1998; Dunkel, 1995; Tomlinson, 1998). In the ITA field, some commentators and textbook writers suggest that ITAs not only need all four skills of "reading, speaking, listening, and writing" for success (Byrd & Constantinides, 1988, p. 124) but that "listening comprehension" should be considered as an area for language development noting that "it is unrealistic to expect global changes in pronunciation after a three week pre-term course, whereas significant improvement in listening comprehension is possible" (Smith et al., 1991, p. 159). Thus there is some support within ESP and ITA education for what might be termed input approaches to language development. Certainly, there is support for an input approach for improving L2 speaking among general language education commentators, including Sajavaara, who noted "more attention should now be paid to reception as a basis for improving production" as "speaking exercises" only aim at correct use of the "machinery of speech production" (1987, p. 58). The current study supports a combined production and input orientation to ITA speaking fluency development.

One theoretical avenue to improving L2 fluency. The position taken in this report is that when input is comprehensible (Doughty, 2003), and when it has features which are noticed and processed by the learner, the input may then potentially change learners' mental representations (Abe, 2009; Skehan, 1998). ITAs (L2 learners) may then be able use these altered mental representations in their speech production (B. VanPatten, personal communication, November 7, 2009; VanPatten, 2003; see also Sajavaara, 1987 on the role of

previous "experience" with language on speaking fluency). Some features acquired through input and noticing may positively influence ITAs' L2 speaking fluency, including pauses which more often correspond to phrase and clause boundaries (Pennington & Ellis, 2000; Simon, 1980).

Repeated reading as a source of comprehensible input. The current study was motivated by unsolicited comments from English learners in Vietnam and the U.S.A. who participated in studies on Repeated Reading (RR), a methodology for increasing reading fluency and comprehension. A pedagogical expression of Automaticity Theory used in L1 and L2 language education contexts, RR works by increasing readers' familiarity with texts through reading a relatively easy 500-word text repeatedly, both silently and with audio support (audio support means a learner silently reads and listens to the text being read aloud at the same time). Each time a text is read, lower-order word recognition processes become automatized, allowing higher-order comprehension processes to be invoked (for reviews of RR research, see Gorsuch & Taguchi, 2009; Taguchi, Sasamoto, & Gorsuch, 2006). Positive effects from RR on reading fluency and comprehension are apparent in the short term, but also in the long term, with learners engaged in regular treatments over time showing steadily faster reading rates (Taguchi, 1997; Taguchi & Gorsuch, 2002) and comprehension the first time they encounter a new, unpracticed text (Gorsuch & Taguchi, 2008).

L2 learners' unsolicited comments mentioned above systematically pointed to input provided by RR as improving their listening skills, specifically their ability to notice stress and intonation being used with specific words and sentences, and the pronunciation of specific known and unknown words. Learners in the studies also systematically noted that during audio-supported readings, they understood the text better because the input allowed them to hear stress, intonation, and pronunciation in meaningful sentence- and discourse-level contexts (see Gorsuch & Taguchi, 2009 for frequencies of comments and for specific examples). It will be useful here to outline a typical RR treatment, which was used for the current study:

1) Participants read an approximately 500-word segment of a popular science text once while timing themselves with a stopwatch. They write their times on a timelog sheet.
2) Participants then read the text a second time and then a third time while listening to it on an audiotape or being read aloud by the teacher. On the third listening/reading, participants are given the option to quietly speak along with the audio model.
3) Participants finally read the text a fourth time, timing themselves for each reading and marking each time on their time log sheet.
4) At the end of the session, participants write a short report either in their L1 or L2, their choice.

It is argued here that RR was a means of making extended streams of input comprehensible in the current study. Note in the procedure above that learners read and heard the input multiple times. Post-treatment written reports and consistently increasing reading rates suggested learners *were* comprehending the input. Because phrase and clause boundaries were salient in this input, it is more than possible that these features, among others, were noticed by ITAs in the 20, twice-a-week, 25-minute sessions (see Olynak, Anglejan, &

Sankoff, 1990, p. 152 for commentary on how texts read aloud constitute "ideal delivery" of English in terms of pauses). These features that were noticed may have been processed, and may have become available for use in output.

Research Questions

In order to establish whether the RR input had at least the potential to be comprehensible, the following research question was posed:

1. Did participants' word-per-minute reading rates and comprehension increase from the 1st to the 20th treatment?

To establish a baseline comparison of participants' fluency and the fluency of native English speaking teaching assistants during teaching sequences, the following question was posed:

2. How do the pre-treatment participants and their native English speaking counterparts compare on teaching talk sequences in terms of rate of speech, percentage of fluent pause groups, split pause groups, false start pause groups, and filler pause groups?

In order to learn whether the 10-week combined input and speaking production approach positively influenced the speaking fluency of ITAs, the following research question was posed:

3. Does participants' spoken fluency change over time? Specifically: rate of speech, and percentage of fluent, split, false start, and filler pause groups?

Further, in order to determine whether participants getting input treatments in addition to conventional production-oriented instruction differed from a control group of ITAs getting only production-oriented instruction, the following question was posed:

4. How does the spoken fluency of the participants differ from an ITA control group on teaching talk sequences in terms of rate of speech, and percentage of fluent, split, false start, and filler pause groups?

Method

Participants

There were three groups of participants. The input ITA group ($n = 28$) were two intact classes of a semester-long ITA preparation course offered at a large southwestern university. They were graduate students in biology, business, chemistry, education, engineering, food processing, personal financial planning, physics, and retailing, and were from Algeria, China, Costa Rica, India, Japan, Korea, Taiwan, Thailand, and Turkey. They were in their mid- to late 20s, and 18 were males and 10 were females. Tests of homogeneity of variance between the two intact classes on all measures were $p < .05$ so the two classes were combined to make one group.

The second group, called native speaker instructors ($n = 4$), were graduate students in biology and chemistry, and were instructors of record for classes and labs. Two were male and two were female. All were American-born, native speakers of English. Their classes were recorded for an unrelated materials development project. The third group, the control ITA group ($n = 10$), were an intact class which attended an ITA preparation course which used

production-oriented instruction and materials, but no input treatments. This course was not taught by the author. Seven were male and three were female, and they were from China, Korea, India, Taiwan, and Vietnam.

Materials

Repeated reading (RR) materials. The RR materials which comprised the input treatments were 350-600 word segments of popular science texts taken from *Science for Kids* (http://www.sciencenewsforkids.org), which includes original texts written at the junior and senior high level on agriculture, computers, plants, and weather, among other things. Some texts were short enough to be stand-alone treatments, while others needed to be segmented at discoursally relevant points and turned into a series of treatments (one segment done on one day, the next on the next day, etc.). As noted above, the ITA input group was combined from two groups who took the ITA preparation course during two different semesters. While the two groups had some RR input texts that were unique to that group, they still used 11 texts (out of 20 total) in common. *T*-tests comparing the length and difficulty estimates of the texts the two sub-groups read that were unique to their sub-group were found to be non-significant at $p < .025$. The combined ITA input group read the following texts in common (see Table 1):

Table 1. Repeated Reading Texts Used for Input in the Combined ITA Input Group

Title	Author	Words	Flesch Kincaid Grade Level
Putting the squeeze on toothpaste A	Sohn (2006a)	613	8.1
Putting the squeeze on toothpaste B		270	6.7
Morphing a wing	Sohn (2007a)	360	7.8
Chocolate rules A	Sohn (2005a)	452	9.1
Chocolate rules B		372	8.0
Listen and learn A	Sohn (2007b)	511	8.7
Listen and learn B		238	7.9
Fungus hunt A	Sohn (2005b)	546	7.6
Fungus hunt B		388	8.9
To hunt a dragonfly A	Sohn (2006b)	374	8.9
To hunt a dragonfly B		524	8.4

The audio-supported part of the input treatments was audio recorded and live readings of the texts by male and female native speakers of English. Readers were advised to read at a relaxed pace with normal intonation and pauses. Readers reported practicing reading the treatment texts aloud at least twice before their "performances," and read on average at 145

words per minute. A transcription from one of the recorded texts (*Morping a Wing*, Sohn, 2007a) shows pause groups demarked by //:

> a silver-colored test wing beneath a jet has a changeable shape // its developers completed flight tests of this new wing last month // "this is something that the aerospace community has been after for a long time" // says aerospace engineer // Peter M. Flick // he works at the Air Force Research Laboratory// at Wright-Patterson Air Force Base // near Dayton Ohio // Flick heads the program that's funding the wing's development // some military jets already have wings that can change shape // the old technology however // is bulky // heavy // and impractical for wide use.

Speaking production-oriented material. At the time of the study, original material from a textbook project was being piloted, and students had the materials in word-processed, photocopied, and bound form. The course included direct instruction on aspects of discourse intonation, including thought groups ("a series of one or more words that form a meaningful and grammatical idea," Meyers, Pickering, Gorsuch, & Griffee, 2010, Chapter 2, p. 1, and referred to as "pause groups" elsewhere in this study), prominence (primary stress), and tone choices ("significant change in pitch movement on the prominent syllable of the focus word" Chapter 4, p. 1), both as stand-alone content, and also in the contexts of introducing oneself and one's courses, leading labs and classes, giving instructions and advice, and asking and answering questions.

Participants heard short audio segments of "good" and "bad" examples of native- and non-native speakers' use of pauses, etc., recorded themselves, and transcribed their recordings for individual and small group critique, read transcribed talks aloud in pairs, and predicted and confirmed where pauses, prominence, etc. should go in a transcript. An example of a transcript participants had to write in // to predict pause groups and then listen to confirm is (Gorsuch et al., 2010, p. 89):

> but real quick today we're looking at freezing point depression // uh experiment one so if you'll all open your lab manuals // you all can read the objective there for youself // but you're supposed to determine uh // you're to investigate the freezing points of pure water

Participants also gave multiple mini-presentations and engaged in self-evaluation activities designed to build metacognitive awareness of weaknesses and strengths in speaking. Both the input ITA (II) and control ITA (CI) groups studied the same materials for the same period of time (14 weeks) and at the same intensity (the same number of chapters were "covered").

Procedure

The input ITA (II) group had RR input treatments for ten weeks, twice a week. They did their RR session as a group, either immediately before or after their regular class meeting. The procedure for each RR treatment is given above in the literature review. Participants had the opportunity to ask questions on pronunciation and word meaning after the first silent reading. Each treatment took 20-25 minutes. Reading time data were recorded by participants on their reading time logs and were tracked as word-per-minute (WPM) rates.

At the beginning and end of the course two pre- and post-test measures were taken. The first was an RR recall pre- and post-test in which participants read stand-alone texts from *Science for Kids* (the same source as the treatment materials): *A New Basketball Gets Slick* (Sohn, 2006c, 471 words, Flesch Kincaid Grade Level = 9.3) for the pre-test and *Glider in the Family* (Sohn, 2007c, 424 words, Flesch Kincaid Grade Level = 9.2) for the post-test. Participants did the pre- and post-test using the RR procedure and were asked to write as much as they could remember after their fourth and final reading.

The second pre- and post-test measure was a seven- to eight-minute presentation on a topic within participants' fields. Sample pre-test presentation topics were: The characteristics of fresh meat, atomic weight, the food chain, risk management, mesoporous materials, electromagnetism, and corporate personality types. Sample post-test presentations were: The structure of bacteria, die casting, Piaget's theory of cognitive development, food rheology, and salmonella. Because speaking performance is greatly influenced by task difficulty (e.g., Pawley & Syder, 2000), participants were asked to self-report on presentation difficulty. Correlational analyses between participants' ratings on pre- and post-task presentation difficulty, and all measures for RQ #3 (fluent pause groups, etc.) revealed no significant relationships. Thus, pre- or post-test performances did not significantly vary as a result of task difficulty. Both pre- and post-test presentations were recorded and transcribed.

Analyses

For RQ #1 on reading rate and comprehension, input ITA (II) participants' first and fourth word-per-minute reading rates for the 1st through the 20th RR session were calculated. For the recall pre- and post-tests, scoring was accomplished by all reasonably complete propositional units being assigned a point. The points were totaled and converted into percentages for total propositional units for each participant. Intrarater reliability for the recall test data was estimated after a six months' hiatus at 95%. To analyze changes in participants' reading rate, a repeated measures, within-subjects 2 X 2 ANOVA was done with Occasion as one variable (the 1st and 20th session) and Sequence as another (1st silent reading rate and 4th silent reading rate). To analyze changes in participants' comprehension, a repeated measures, within subjects one-way ANOVA was done comparing the percentage of text propositions recalled after the fourth reading of a pre-test text, and then after the fourth reading of a post-test text (variable = Occasion).

Prior to any further analyses, all II group recordings and transcripts were coded to ensure the rater (the researcher) would not know whether she was analyzing a pre- or post-test performance, nor whose individual performance she was analyzing. First, portions of transcripts for the II group (n = 28) were selected for analysis. In general, these portions comprised the main body of their pre- and post-test presentations where they spoke continuously for approximately four to five minutes. Second, the recordings were played and pauses of .02 seconds or greater were marked. Third, pause groups were categorized as "fluent," "split," "false start," or "filler" (see Literature Review above). Fourth, the recordings were listened to a final time to confirm pauses and categorizations. To control for differences in the length of the presentations, numbers of "fluent," "split," etc. pause groups

were converted to percentages of the total number of pause groups. Finally, rate of speech was estimated.

Because there was so much data per participant (up to hundreds of pause groups, for example), and because there were 65 total recordings and transcripts, and because listening to each recording was so time consuming, it was not possible to find an independent rater to establish interrater reliability. To establish consistency of results, *intrarater* reliability checks were done on pause markings, and pause group categorizations between six to eight months after the initial analyses were done for input ITA (II), control ITA (CI), and native speaker instructors (NSI) groups. There were few discrepancies in the second ratings.

For RQ #2, the II group's pre-test transcripts were compared to four- to five-minute sequences of continuous talk class recordings and transcriptions of the NSI group ($n = 4$). Five *t*-tests were done with $p = .02$ (.10 divided by 5 for five comparisons) for speech rate, and percentages of fluent, split, false start, and filler pause groups.

For RQ #3, a repeated measures, within-subjects ANOVA was done comparing II group's rate of speech from the pre- to the post-test with p set at .05. A repeated measures, within-group 2 X 4 ANOVA was done with Occasion as one variable (pre- and post-test) and Pause Group Type as the second variable (fluent, split, false start, and filler).

Finally, for RQ #4, five to six minute portions of the CI group's presentations ($n = 10$) taken at the end of their course were transcribed and analyzed using the same procedure described above for the II group. Because pre-course recordings for the CI group were not available, and because pre-treatment equivalence between the groups had to be established in some way, the pre-course, pre-treatment SPEAK test scores of all II and CI group members who had documented pre-course scores were compared (the SPEAK test is the institutional version of the Test of Spoken English, which "measures the ability of nonnative speakers of English to communicate effectively" (Educational Testing Service, 2009)). While all ten CI group members had documented pre-course SPEAK scores, only 17 II group members had them. This partial II group mean for the pre-treatment SPEAK was $M = 41.47$, with $SD = 2.93$. The CI group mean for the pre-course SPEAK was somewhat higher at $M = 44.00$, $SD = 3.16$. This difference was significant with a two-tailed *t*-test ($t = -2.01$, $p = .046$). Thus, comparisons for RQ #4 described below were done with partial II $n = 17$ and CI $n = 10$. The partial II (pII) group's post-test speech rate was compared to the CI group's speech rate using a between-subjects ANOVA with Group as the independent variable, speech rate as the dependent variable, and pre-course, pre-treatment SPEAK as a covariate. The two groups' pause group types (fluent, split, false start, and filler) were compared using a between-subjects MANOVA with Group as the independent variable (pII versus CI), and pause group type as the dependent variables, and pre-course, pre-treatment SPEAK scores as a covariate.

Results

The descriptive statistics for the word-per-minute reading rate portion of RQ #1 are seen in Table 2:

Table 2. Descriptive Statistics for WPM Reading Rate for 1st and 4th Readings for the 1st and 20th RR Sessions (n = 28)

		M	SD
First session	1st silent reading	134.09	29.11
	4th silent reading	171.33	36.16
Twentieth session	1st silent reading	160.86	56.41
	4th silent reading	198.83	64.18

The repeated measures ANOVA was statistically significant for both variables, Occasion (1st to the 20th sessions) at $p = .005$, effect size $eta^2 = .253$; and for Sequence (1st to the 4th silent reading) at $p = .000$, effect size $eta^2 = .569$. The input ITA group read faster not only within RR sessions ($M = 134.09$ WPM for the 1st reading of the 1st session versus 171.22 WPM for the 4th reading, and 160.86 WPM for the 1st reading of the 20th session versus 198.83 WPM for the 4th reading), but also in the long term from the beginning to the end of the ten-week treatments. Note that participants read a new, unpracticed text faster in the 20th session (134.09 versus 160.86 WPM).

The descriptive statistics for the comprehension portion of RQ #1 are as follows: On average, II participants recalled 16.19% of all propositions ($SD = 7.89\%$) of the pre-test text on their 4th reading, and eleven weeks later recalled on average 31.54% of all propositions ($SD = 13.13\%$) on the post-test text on their 4th reading. This increase was statistically significant at $p = .000$, effect size $eta^2 = .576$.

For RQ #2, the pre-test presentation performances of the input ITA (II) group was compared to the native speaker instructor (NSI) group on five measures. The descriptive statistics are in Table 3:

Table 3. Input ITA (N = 28) and Native Speaker Instructor (N = 4) Group Comparisons

	Input ITA Group		**NS Instructor Group**	
	M	SD	M	SD
Speech rate (WPM)	87.35	17.14	126.14	10.15
Percent of fluent pause groups	43.34	11.39	73.41	10.42
Percent of split pause groups	37.88	12.56	13.88	5.81
Percent of false start pause groups	8.11	5.27	1.57	1.19
Percent of filler pause groups	10.69	5.69	8.97	7.41

Table 3 shows that the II group and NSI group differed significantly on three of the five measures. The NSI group spoke at a faster rate than the NSI group (M = 126.14 WPM versus 87.35 WPM; note also the smaller SD of 10.15 for the NSI group, suggesting less variation on this measure). This difference was statistically significant at p = .000, effect size eta^2 = .39. The NSI group also used a higher percentage of fluent pause groups (M = 73.41 versus M = 43.43), a difference which was significant at p = .000, effect size eta^2 = .38. In contrast, the II group used a higher percentage of split pause groups (M = 37.88 versus M = 13.88) which was significant at p = .001, effect size eta^2 = .41. The II group also used higher percentage of false start pause groups (M = 8.11 versus M = 1.57), but this difference was not significant at P = .021, effect size eta^2 = .16. Finally, the II group used a somewhat higher percentage of filler pause groups than the NSI group (M = 10.69 versus M = 8.97), although this difference was not significant, and the effect size eta^2 was less than 1%.

The descriptive statistics for the pre- and post-test speaking rate, and percentage of fluent, split, false start, and filler pause groups for RQ #3 are seen in Table 4:

Table 4. Pre- and Post-Test Descriptive Statistics for Input ITA Group (n = 28)

	Pre-test		Post-test	
	M	SD	M	SD
Speech rate (WPM)	87.35	17.14	91.51	10.77
Fluent pause groups (%)	43.34	11.38	48.36	11.5
Split pause groups (%)	37.88	12.55	32.00	10.69
False start pause groups (%)	8.11	5.26	6.33	3.85
Filler pause groups (%)	10.69	5.69	11.11	5.19

There was no significant change in the input ITAs' speech rate (effect size eta^2 = .064), although the mean rose slightly for the post-test, and the SD narrowed suggesting less variation (M = 87.35 and SD = 17.14 versus M = 91.51 and SD = 10.77). However, the IIs used a higher percentage of fluent pause groups (M = 48.36 for the post-test versus M = 43.43 for the pre-test). Input ITAs used fewer split and false start pause groups in their post-tests, decreasing from M = 37.88 to M = 32.00 for split pause groups and from M = 8.11 to M = 6.33 for false start pause groups. The SDs for both split and false start pause groups narrowed, suggesting less variation in input ITAs' performances. The II group's use of filler pause groups used remained about the same (M = 10.66 for the pre-test versus M = 11.11 for the post-test).

The variable of Occasion (pre-test versus post-test) was statistically significant at p = .025 (eta^2 = .173) suggesting that the II group's use of different pause groups changed over time. Not surprisingly, the variable of pause group type was significant at p = .000, effect size eta^2

= .971 (no one would expect speakers to use 25% each of the four pause group types in any natural performance). Finally, there was a significant Occasion and Pause Group Type interaction (p = .006, effect size eta^2 = .388) suggesting that the other two pause group types, fluent and split, did change in terms of proportion of all pause groups used over time, which can be seen in Table 4 above.

The descriptive statistics for the partial Input ITA (pII) and Control ITA (CI) group's speaking rate, and percentage of fluent, split, false start, and filler pause groups for RQ #4 are seen in Table 5:

Table 5. Partial input ITA (n = 17) and Control ITA (n = 10) Comparisons

	Partial Input ITA		Control ITA	
	M	SD	M	SD
Speech rate (WPM)	90.47	10.41	97.56	24.97
Fluent pause groups (%)	46.97	12.66	48.32	8.24
Split pause groups (%)	33.91	10.87	38.16	11.55
False start pause groups (%)	7.17	3.36	3.13	2.33
Filler pause groups (%)	11.93	4.47	10.39	5.08

It must be clarified that the pII group results reported in Table 5 above are from only 17 of the total 28 members of the group, due to the fact that only 17 of the II group had taken the SPEAK test prior to the treatment, whereas all 10 members of the CI group had. In order to use pre-treatment, pre-course SPEAK test scores as a covariate (and thus ensuring a statistically equal playing field in terms of statistical significance and effect size) the n size of the II group had to be reduced by 39%. Note then that the descriptive statistics for the full II group (n = 28) reported in Table 4 above are slightly different than the partial II group (n = 17) reported in Table 5 above. For instance, the II group used a higher percentage of fluent pause groups (M = 48.36) while the pII group used fewer (M = 46.97). These differences, caused by a sampling artefact, must be taken into account when viewing the statistical analyses.

At the end of the treatment, the pII group spoke at a somewhat slower rate than the CI group at the end of their course (M = 90.47 WPM versus M = 97.56). This difference was not significant at p = .466. The pII group (M = 46.97) did not use quite as many fluent pause groups as the CI group did (M = 48.32). At the same time, the pII group used *fewer* split pause groups than the CI group (M = 33.91 versus M = 38.16). The pII group used more false start pause groups than the CI group (M = 7.17 versus M = 3.13), and the two groups used about the same number of filler pause groups (pII M = 11.93, CI M = 10.39). The main effect for Group approached significance but did not reach a critical value (p = .066, effect size eta^2 = .331), meaning that while not statistically significant, the Group variable effect size was moderate and still accounted for 33.1% of the variance in the analysis model. Between-subject effects for fluent, split, and filler pause groups were not significant. The between-

subjects effect for false start pause groups was significant at $p = .003$, effect size $eta^2 = .314$, reflecting the pII group using more false start pause groups at end of their treatment + course, than the CI group after their course alone.

Discussion

Universal needs of ITAs as language learners. The overall purpose of this study was to propose a greater role for input in ITA educators' enduring quest to improve ITAs' spoken fluency. Another purpose was to propose that the language learning challenges of ITAs be viewed as embedded in second language acquisition theories, which stipulate a need for input, noticing, and output. Thus, regardless of institutions or disciplines in which ITAs use English for teaching, ITAs' language learning needs and requirements are universal. All L2 learners, including ITAs, need comprehensible input, and need to notice formal aspects of the input. The current study highlighted empirical evidence that suggests a mixed input and output orientation to fluency instruction succeeded in increasing the percentage of fluent (grammatically intact) pause groups used by ITAs in extemporaneous talk, and also in decreasing the percentage of split pause groups in which phrasal boundaries are violated by pauses (see Table 4 above). When compared to a control group which had only a production oriented course, the partial ($n = 17$) input ITA group used fewer split pause groups, and more false start groups (see Table 5 above). It is important to note that false start pause groups, while distracting with their self-repairs and on-line revisions, often represent syntactically intact phrases and clauses. Note, however, that in the full $n = 28$ II group that they used about the same percentage of false start pause groups in the pre-test as they did in the post-test (see Table 4). The difference in false start pause group use between the $n = 17$ II group and CI group may have pre-existed the treatments + course for the II group and the course for the CI group. False start pause groups may simply be one aspect of temporal fluency, as it is measured in the current study, that was not as sensitive to the input treatments as fluent and split pause groups.

Fluency and ITA program curricula. In Table 3 above, the differences between ITAs and native-English speaking TAs for temporal aspects of spoken English is clear. The ITAs used fewer fluent pause groups, and more split pause groups than the American TAs (see Table 3 above). This information is not highlighted to suggest that ITAs be viewed in terms of "deficits" in comparison to American TAs but rather to suggest that these aspects of spoken fluency are indicative of language development issues which are salient to ITA education curricula.

Basically, pauses which "break up the timing pattern" of spoken English (split pause groups) are "tentative" (Butterworth, 1980, p. 158) and suggests that learners are not able to plan beyond sentence level grammar (Deese, 1980). If ITAs are struggling with encoding sentence level grammar, they cannot plan at the discourse level (Deese, 1980), something arguably needed for teaching (for instance, how can an ITA plan the content and timing of examples in their talk if they cannot plan at the discourse level?). Ejzenberg (2000) provides a description of the cognitive processes involved in low and high levels of fluency. Fluent speakers have longer "fluent units" and a greater repertoire of formulaic speech chunks which "are apparently stored, retrieved, and ultimately uttered as one indivisible or preassembled unit" (p. 305). Less fluent speakers also "have" speech chunks but "fail to retrieve fixed and semi-

fixed chunks in grammatically accurate ways" (p. 307), which may partially account for split pause groups. Thus ITAs not only need help with building automaticity in encoding with speech production (through practice and metacognitive awareness building), but also with the L2 in their minds, the very stuff they need to encode to begin with (through more input). The point here is that building L2 fluency is a developmental cognitive process, and takes much time and concerted effort by learners and teachers to develop. Moderate to intensive input treatments may facilitate this process. Explicit fluency building deserves and requires expanded and consistent treatment in ITA program curricula.

One intriguing detail about the input ITA group's speech rate over time (see Table 4 above) also points to the importance of an input approach in addition to production-oriented instruction to improve spoken fluency. With a production-only approach focused on building metacognitive awareness of and control over fluent speech performance, it may be that ITAs would have to slow down their speech rate in order to monitor and perform. In the current study, participants' speech rates increased slightly *and* their performance was more accurate (a higher percentage of fluent pause groups used). It is possible that the input gave participants more to work with in terms of their cognitive linguistic resources.

Audio-supported repeated reading as one means of input for spoken fluency development. It is argued here that the audio-supported RR treatments provided the amount and type of input that the input ITA group needed to propel their fluency development (see Table 4 above). Participants comprehended the passages after repeated experience with them (see Table 2 and the Results section above). As their engagement with word-level comprehension processes became automatized, participants freed up enough of their cognitive resources to invoke higher order comprehension processes, and I would argue, to attend to salient features in the audio model of the text. Pauses were salient in the input, likely made so by the pauses themselves, and also by the intonation contours which corresponded to pause group and sentence boundaries (see Butterworth, 1980, p. 168, on the "phonemic clause"). Pennington and Ellis (2000, p. 372) posit that such intonation contours, as prosody, form "part of the memory representation which listeners form of the input." While detailed discussion of this important aspect of audio-supported RR as input must be relegated to another, related study, suffice to say this is an area that needs further research, and may be one of the main reasons participants may have been able to form new mental representations, and make use of them in their post-test presentations.

Suggestions on implementing an input approach. Suggestions are: 1. Maintain at least a moderate-intensity input program; 2. Do the input sessions as a face-to-face activity; 3. Work with combined visual and audio inputs; 4. Select texts that are relevant to ITAs, and somewhat easy.

In the institution where this study took place, ITA preparation classes meet only twice a week for 80 minutes each. Class sizes average 16. In order to maintain an input program of moderate intensity, two 25-minute sessions per week were required. It was found that participants could, and would, attend a session either immediately before or after their regular class meeting. For an intensive workshop of ten or more days, 25 minutes per day is not difficult to do.

It will be tempting for many ITAs and ITA educators to work with input treatments online asynchronously, or in some other non-face-to-face format. I do not recommend this. In order for input to be comprehensible, the input must be "met" repeated times, and ITAs need to be able to ask questions about word pronunciation or meaning in the texts, many of which cannot be anticipated. Many ITAs, left to their own devices, will listen or read to texts only once, which will result in only a sparse comprehension of the input. It is unlikely they would ever get to the point where enough of their attentional resources were free to notice pause group boundaries, prosodic contours, or whatever else instructors think they ought to notice. How the brain deals with L2 input is not obvious to those outside the field, and this is one case where ITA educators, as specialists, need to insist on methods which will ensure the input is comprehensible. One other positive aspect of face-to-face treatments is the opportunity for ITAs to track their own progress on a treatment-by-treatment basis. In the institution where this study took place, ITAs timed their own reading with stop watches, and then kept track on a log sheet. They could see their reading times go down over time, and often remarked on this while at the same time noting they understood the texts more.

It is also suggested that combined visual and audio input treatments be used. On one hand, the input materials must contain language and concepts that are academic in order to be relevant to ITAs' disciplines (see below). This implies a text, or a recorded classroom talk, which would contain condensed narrative and expository elements, as well as academic vocabulary. On the other hand, the input must be comprehensible. If ITAs were only to hear a text or classroom talk, it seems less likely they will comprehend it, even after repeated exposures. Bassetti (2009, p. 191), in arguing for a greater understanding of the role of written input for second language phonology acquisition, notes that written "representations provide a visual analysis of language." ITAs may need to see and hear input to comprehend and to notice salient features of temporal fluency.

Finally, it is important to select texts for input treatments that are relevant to ITAs, and are fairly easy. Relevance not only has to do with topic but also with communicative functions ITAs may have to invoke while teaching, such as using examples, underscoring important information, explaining a process, or giving definitions. With some searching, sources for such texts can be found. One source is http://www.sciencenewsforkids.org. Another source, recently reported by an ITA educator is http://www.commoncraft.com, which uses "plain English" to explain topics such as technology and society (C. Quarterman, ITA listserv communication, March 15, 2010). Ensuring the input is comprehensible is key. This means using texts which are fairly easy, and have no more than 500 words. While the Flesch Kincaid Grade Level check (available under "reading statistics" in many word processing programs) is only rough at best, it is effective for the general selecting and sequencing of texts. Post-treatment reports from participants in the current study suggested that texts at the 7th to 8th grade level with no more than five unknown words were comprehended best, and generated the most post-treatment comments which exhibited higher order comprehension and noticing of linguistic features.

Study limitations and future areas of research. This study has two potential limitations. The main limitation is that causality for the input treatments is difficult to establish. Comparisons with the partial (n = 17) II and CI groups suggested subtly different post-treatment, post-course fluency profiles (see Table 5 above). Yet it will never be clear whether the

comparison was valid, as both the partial input ITA and control groups were samples of convenience, defined by class membership, and whether documented pre-treatment and pre-course SPEAK test scores were available. In other words, it will never be known whether pII or CI group participants were drawn from the same population, and whether even the full (*n* = 28) II and CI groups are comparable. Nonetheless it is important to point out that whether the ITA input group improved due to the production-oriented instruction, the input treatments, or a combination of the two, the spoken fluency of the II group did improve (see Table 4 above) and in somewhat different ways than the CI group (Table 5). More importantly, there are compelling theoretical reasons to believe that comprehensible input and attention to features in that input is necessary for language acquisition. Arguably, these elements were provided in the treatments.

Another limitation is the difficulty in determining the reliability of the researcher judgments of where pauses occurred, and whether a pause group ought to be judged fluent, split, or false start. Intrarater reliability procedures were done, but some readers may still doubt the data reliability. My only answer is to point out again that the audio files of the input ITA group were coded and randomized, and it was not possible to know whether it was a pre- or a post-test being analyzed. Further, the control group was an intact class not chosen by me; rather, it was "selected" by impersonal processes of course registration. The same intrarater reliability procedures were followed, as with the input ITA group. Experience with this project reveals the real difficulty in getting *interrater* reliability on complex human performances such as the features of spoken L2 temporal fluency in extended discourse.

Areas of future research include empirically exploring ways to more reliably direct ITAs' attention to fluency features in input. This is a formidable challenge when dealing with extended, authentic discourse, and when there are so many features of temporal fluency that can be focused on. Chun, Hardison, & Pennington (2008) reviewed multiple studies where L2 learners were trained to perceive and produce different, prespecified features of English prosody using various methodologies with single sentences. Hardison (2005, p. 179) created computer-based training materials by selecting "problematic" "sequences" from 28 learners' own extemporaneous presentations and turning them into individualized two-week training programs with contour lines on a computer screen. She found positive effects on external judges' ratings of participants' post-training speech, but it is not clear which specific areas of prosody each individual participant worked on, nor how improvement in these areas may have contributed to judges' perceptions of global improvement.

In the current study, it can be argued that participants' attention was focused in a general way by being engaged with the production-focused instruction, which featured explicit instruction on pause groups, prominence, and tone choices. Yet there seems to be something else going on, where the participants are working out the relationship between pauses, intonation contours, and phrase and clause and sentence boundaries through the RR input. This knowledge was apparently available for use in participants' extemporaneous talk. Perhaps this language development should be the real focus of further research, with an eye to more explicitly exploring the relationship between extended comprehensible input, changes in learners' ability to connect the L2 sound system to L2 they have only studied visually in the past, and changes in learners' spoken fluency. Perhaps there are systematic ways to further

explore what aspects of the input ITAs are attending to, and whether and how ITAs may be using the input differently due to ability level, or L1.

Conclusion

This study addressed the issue of whether an input approach had a place in spoken fluency development for international teaching assistants, an L2 learner population facing multiple, intractable challenges in using their L2s for professional purposes. The results suggested that a mixed input and production oriented approach helped participants improve their L2 fluency in extemporaneous teaching presentations. Further explorations of input approaches to ITA education are warranted, both for practical and theoretical reasons. Finally, ITAs should be viewed first and foremost as language learners, whose learning processes are the same as any other L2 learner. Thus, the frequency, length, and curricula of ITA classes and workshops should be decided on the basis of ITAs' language learning needs. Spoken fluency development, and the language development that underlies this, takes time, and concerted and persistent effort.

About the Author

Greta Gorsuch (Temple University Japan, Ed.D.) is an associate professor of applied linguistics and second language studies at Texas Tech University. She is interested in international teaching assistant education, domain and local teacher theories in language learning, and EFL education overseas. She is currently researching reading fluency, the learning of English prosody, and concept development in graduate students.

References

Abe, H. (2009). The effect of interactive input enhancement on the acquisition of the English connected speech by Japanese college students. *Phonetics Teaching & Learning Proceedings, University College London*. Retrieved February 18, 2010, from http://www.phon.ucl.ac.uk/ptlc/proceedings/ptlcpaper_29e.pdf.

Alsberg, J. (2002). Effecting change in pronunciation: Teaching ITAs to teach themselves. In W. Davis, J. Smith, & R. Smith (Eds.). *Ready to teach: Graduate teaching assistants prepare for today and tomorrow* (pp. 139-146). Stillwater, OK: New Forums Press, Inc.

Ard, J. (1987). The foreign TA problem from an acquisition-theoretic point of view. *English for Specific Purposes, 6,* 133-144.

Asker, B. (1998). Student reticence and oral testing: A Hong Kong study of willingness to communicate. *Communication Research Reports*, *15*(2), 162-169.

Bassetti, B. (2009). Orthographic input and second language phonology. In T. Piske & M. Young-Scholten (Eds.). *Input matters in SLA* (pp. 191-206). Bristol, UK: Multilingual Matters.

Butcher, A. (1980). Pause and syntactic structure. In H. Dechert & M. Raupach (Eds.). *Temporal variables in speech* (pp. 85-90). The Hague: Mouton Publishers.

Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (Ed.). *Language production: Speech and talk* (pp. 155-176). London: Academic Press.

Byrd, P. & Constantinides, J. (1988). FTA training programs: Searching for appropriate teaching styles. *English for Specific Purposes, 7*, 123-129.

Byrd, P., & Constantinides, J. (1995). Textbooks for training programs for international teaching assistants at U.S. universities. In P. Byrd (Ed.). *Material writer's guide* (pp. 137-146). Boston, MA: Heinle & Heinle Publishers.

Byrd, P., Constantinides, J., & Pennington, M. (1989). *The foreign teaching assistant's manual*. New York, NY: Collier MacMillan.

Chesser, V. Carroll, S., Macero, J., & Tice, S. (2002). Oral communication in teaching: A program assessment. In W. Davis, J. Smith, & R. Smith (Eds.). *Ready to teach: Graduate teaching assistants prepare for today and tomorrow* (pp. 151-158). Stillwater, OK: New Forums Press, Inc.

Chun, D., Hardison, D., & Pennington, M. (2008). Technologies for prosody in context. In J. Hansen Edwards & M. Zampini (Eds.). *Phonology and second language acquisition* (pp. 323-346). Amsterdam: John Benjamins Publishing Company.

Cornell University (2007). ITA courses. Retrieved June, 2009 from http://www.cte.cornell.edu/campus/itadp/.

Crystal, D. & Davy, D. (1969). *Investigating English style*. Bloomington, IN: Indiana University Press.

Deese, J. (1980). Pauses, prosody, and the demands of production in language. In H. Dechert & M. Raupach (Eds.). *Temporal variables in speech* (pp. 69-84). The Hague: Mouton Publishers.

Dickerson, W. (2002). Grading pronunciation progress: the roles of absolute and relative scales. In W. Davis, J. Smith, & R. Smith (Eds.). *Ready to teach: Graduate teaching assistants prepare for today and tomorrow* (pp. 159-166). Stillwater, OK: New Forums Press, Inc.

Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughty & M. Long (Eds.). *The handbook of second language acquisition* (pp. 256-310). Malden, MA: Blackwell Publishing Ltd.

Dudley-Evans, T. & St. John, M.J. (1998). *Developments in English for specific purposes*. Cambridge: Cambridge University Press.

Dunkel, P. (1995). Authentic second/foreign language listening texts: Issues of definition, operationalization, and application. In P. Byrd (Ed.). *Material writer's guide* (pp. 95-106). Boston, MA: Heinle & Heinle Publishers.

Educational Testing Service (2009). *TSE-Test of spoken English*. Retrieved April 7, 2010 at: http://www.ets.org/portal/site/ets/menuitem.fab2360b1645a1de9b3a0779f1751509/?vgnextoid=b5d7d898c84f4010VgnVCM10000022f95190RCRD.

Ejzenberg, R. (2000). The juggling act of oral fluency: A psycho-sociolinguistic metaphor. In H. Riggenbach (Ed.). *Perspectives on fluency* (pp. 287-313). Ann Arbor, MI: The University of Michigan Press.

Ellis, R. (1990). Individual learning styles in classroom second language development. In J. de Jong & D. Stevenson (Eds.). *Individualizing the assessment of language abilities* (pp. 83-96). Clevedon, England: Multilingual Matters Ltd.

Ellis, R. (1997). *SLA research and language teaching*. Oxford: Oxford University Press.

Ford, J., Gappa, L., Wendorff, J., & Wright, D. (1991). Model of an ITA institute. In J. Nyquist, R. Abbott, D. Wulff, & J. Sprague (Eds.). *Preparing the professoriate of tomorrow to teach*. (pp. 341-349). Dubuque, IA: Kendall Hunt.

Goldman Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. London: Academic Press.

Gorsuch, G.J. (2003). The educational cultures of international teaching assistants and U.S. universities. *TESL-EJ, 7*(3), 1-17 [On-line]. Available: http://www.tesl-ej.org/wordpress/issues/volume7/ej27/ej27a1/.

Gorsuch, G.J. (2006a). Classic challenges in ITA assessment. In D. Kaufman & B. Brownworth (Eds.). *Professional development of international teaching assistants* (pp. 69-80). Alexandria, VA: TESOL.

Gorsuch, G.J. (2006b). Discipline-specific practica for international teaching assistants. *English for Specific Purposes, 25*, 90-108.

Gorsuch, G.J. (2008). Adults becoming bilinguals and becoming teachers: Pushing outside the comfort zone. In P. Dheram (Ed.). *Negotiating empowerment: Studies in English language education* (pp. 159-174). New Delhi, India: Orient Longman Publishers.

Gorsuch, G., Meyers, C.M., Pickering, L., Griffee, D.T. (2010). *English communication for international teaching assistants*. Long Grove, IL: Waveland Press, Inc.

Gorsuch, G.J., Stevens, K., & Brouillette, S. (2003). Collaborative curriculum design for an international teaching assistant workshop. *The Journal of Graduate Teaching Assistant Development, 9*(2), 57-68.

Gorsuch, G.J. & Taguchi, E. (2008). Repeated reading for developing reading fluency and reading comprehension: The case of EFL learners in Vietnam. *System, 36*(2), 253-278.

Gorsuch, G.J. & Taguchi, E. (2009). Repeated reading and its role in an extensive reading program. In A. Cirocki (Ed.) *Extensive reading in English language teaching* (pp. 249-271). Munich, Germany: Lincom.Europa.

Griffee, D., Gorsuch, G.J, Britton, D. & Clardy, C. (2009). Intensive second language instruction for international teaching assistants: How much is effective? In G. Ollington (Ed.). *Teachers and teaching: Strategies, innovations and problem solving* (pp.187-205). Hauppage, NY: Nova Science Publishers, Inc.

Hardison, D. (2005). Contextualized computer-based L2 prosody training: Evaluating the effects of discourse context and video input. *CALICO Journal, 22*(2), 175-190.

Heidish, P.A. (2006). Creating partnerships: International teaching assistant links in a campus-wide chain–the Carnegie Mellon experience. In D. Kaufman & B. Brownworth (Eds.). *Professional development of international teaching assistants* (pp. 165-180). Alexandria, VA: Teachers of English to Speakers of Others Languages, Inc.

Hinofotis, F., Bailey, K., & Stern, S. (1981). Assessing the oral proficiency of prospective foreign teaching assistants: Instrument development. In A. Palmer, P. Groot, & G. Trosper (Eds.). *The construct validation of tests of communicative competence* (pp. 106-126). Washington, D.C.: TESOL, Inc.

Hoekje, B., & Linnell, K. (1994). "Authenticity" in language testing: Evaluating spoken language tests for international teaching assistants. *TESOL Quarterly, 28*(1), 103-126.

International TA Program (2009). International TA program. University of Minnesota. Retrieved March, 2009 from http://www1.umn.edu/ohr/teachlearn/graduate/itap/.

Kopenen, M., & Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In H. Riggenbach (Ed.). *Perspectives on fluency* (pp. 5-24). Ann Arbor, MI: The University of Michigan Press.

Krashen, S. (1985). *The input hypothesis*. London: Longman.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.). *Perspectives on fluency* (pp. 25-42). Ann Arbor, MI: The University of Michigan Press.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Madden, M., & Myers, C. (1994). Introduction. In M. Madden & C. Myers (Eds.). *Discourse and performance of international teaching assistants* (pp. 1-7). Alexandria, VA: Teachers of English to Speakers of Other Languages.

Meyers, C.M., Pickering, L., Gorsuch, G., & Griffee, D.T. (2010). *English communication for international teaching assistant teacher's manual*. Long Grove, IL: Waveland Press, Inc.

Olynak, M., 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, E. Andersen, & S. Krashen (Eds.). *Developing communicative competence in a second language* (pp. 139-155). Boston, MA: Heinle & Heinle Publishers.

Oppenheim, N. (2000). The importance of recurrent sequences for nonnative speaker fluency and cognition. In H. Riggenbach (Ed.). *Perspectives on fluency* (pp. 220-240). Ann Arbor, MI: The University of Michigan Press.

Panvini, D. (2002). Building professional skills: The intersection of ITAs and undergraduates. In W. Davis, J. Smith, & R. Smith (Eds.). *Ready to teach: Graduate teaching assistants prepare for today and tomorrow* (pp. 183-192). Stillwater, OK: New Forums Press, Inc.

Papajohn, D., Alsberg, J., Bair, B., & Willenborg, B. (2002). An ESP program for international teaching assistants. In T. Orr (Ed.). *English for specific purposes* (pp. 89-101). Alexandria, VA: Teachers of English to Speakers of Other Languages, Inc.

Pawley, A. & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J.C. Richards & R. Schmidt (Eds.). *Language and communication* (pp. 191-226). London: Longman.

Pawley, A. & Syder, F. (2000). The one-clause-at-a-time hypothesis. In H. Riggenbach (Ed.). *Perspectives on fluency* (pp. 163-199). Ann Arbor, MI: The University of Michigan Press.

Pennington, M. & Ellis, N. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *The Modern Language Journal, 84*(3), 372-389.

Petro, A. (2006). Addressing the cultural and linguistic needs of students. In D. Kaufman & B. Brownworth (Eds.). *Professional development of international teaching assistants* (pp. 151-163). Alexandria, VA: Teachers of English to Speakers of Others Languages, Inc.

Pialorsi, F. (1984). Toward an anthropology of the classroom: An essay on foreign teaching assistants and U.S. students. In K. Bailey, F. Pialorsi, & J. Zukowski-Faust (Eds.). *Foreign teaching assistants in U.S. universities* (pp. 16-21). Washington, D.C.: National Association for Foreign Student Affairs.

Riggenbach, H. (1989). *Nonnative fluency in dialogue versus monologue speech: A microanalytic approach*. Unpublished doctoral dissertation. University of California at Los Angeles.

Sajavaara, K. (1987). Second language speech production: Factors affecting fluency. In H. Dechert & M. Raupach (Eds.). *Psycholinguistic models of production* (pp. 45-65). Norwood, NJ: Ablex Publishing Corporation.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206-226.

Segalowitz, N. (2000). Automaticity and attentional skill in fluent performance. In H. Riggenbach (Ed.). *Perspectives on fluency* (pp. 200-219). Ann Arbor, MI: The University of Michigan Press.

Simon, H. (1980). How to win at twenty questions with nature. In R. Cole (Ed.). *Perception and production of fluent speech* (pp. 535-548). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Skehan, P. (1998). *A cognitive approach to languge learning*. Oxford: Oxford University Press.

Smith, J., Meyers, C., & Burkhalter, A. (1992). *Communicate: Strategies for international teaching assistants*. Long Grove, IL: Waveland Press, Inc.

Smith, R., Byrd, P., Constantinides, J. & Barrett, R. (1991). Instructional development programs for international TAs: A systems analysis approach. *To Improve the Academy*, 151-168.

Smith, R., Byrd, P., Nelson, G., Barrett, R., Constantinides, J. (1992). *Crossing pedagogical oceans: International teaching assistants in U.S. undergraduate education*. Washington, D.C.: ERIC Clearinghouse on Higher Education.

Sohn, E. (2005a). Chocolate rules. *Science news*. Retrieved February 27, 2007 Available at: http://www.sciencenewsforkids.org/articles/20050209/Feature1.asp.

Sohn, E. (2005b). Fungus hunt. *Science news*. Retrieved February 23, 2007. Available: http://www.sciencenewsforkids.org/articles/20050713/Feature1.asp.

Sohn, E. (2006a). Putting the squeeze on toothpaste. *Science news*. Retrieved May 24, 2006. Available: http://www.sciencenewsforkids.org/articles/20061129/Feature1.asp.

Sohn, E. (2006b). To hunt a dragonfly. *Science news*. Retrieved May 24, 2006. Available: http://www.sciencenewsforkids.org/articles/20061213/Feature1.asp.

Sohn, E. (2006c). A new basketball gets slick. *Science news*. Retrieved May 24, 2006. Available: http://www.sciencenewsforkids.org/articles/20061108/Note3.asp.

Sohn, E. (2007a). Morphing a wing. *Science news*. Retrieved February 21, 2007. Available: http://www.sciencenewsforkids.org/articles/20070103/Note3.asp.

Sohn, E. (2007b). Listen and learn. *Science news*. Retrieved November 15, 2007. Available: http://www.sciencenewsforkids.org/articles/20070926/Feature1.asp.

Sohn, E. (2007c). Gliders in the family. *Science news*. Retrieved November 15, 2007. Available: http://www.sciencenewsforkids.org/articles/20071107/Note2.asp.

Taguchi, E. (1997). The effects of repeated readings on the development of lower identification skills of FL readers. *Reading in a Foreign Language, 11*, 97-119.

Taguchi, E. & Gorsuch, G. (2002). Transfer effects of repeated EFL reading on reading new passages: Silent reading rate and comprehension. *Reading in a Foreign Language, 14*(1), 1-18.

Taguchi, E., Sasamoto, E., & Gorsuch, G. (2006). Developing second and foreign language reading fluency and its effect on comprehension: A missing link. *Reading Matrix, 6*(2), 1-18. http://www.readingmatrix/archives/archives_vol6_no2.html.

Tanner, M. (1991). Incorporating research on question asking into ITA training. In J. Nyquist, R.D. Abbott, D.H. Wulff, & J. Sprague (Eds.). *Preparing the professoriate of tomorrow to teach* (pp. 375-381). Dubuque, IA: Kendall/Hunt Publishing Company.

Thomas, C., & Monoson, P. (1991). Issues related to state-mandates English language proficiency requirements. In J. Nyquist, R.D. Abbott, D.H. Wulff, & J. Sprague (Eds.). *Preparing the professoriate of tomorrow to teach* (pp. 382-392). Dubuque, IA: Kendall/Hunt Publishing Company.

Tomlinson, B. (1998). Introduction. In B. Tomlinson (Ed.). *Materials development in language teaching* (pp. 1-24). Cambridge: Cambridge University Press.

VanPatten, B. (2003). *From input to output*. Boston, MA: McGraw Hill.

Wennerstrom, A. (2000). The role of intonation in second language fluency. In H. Riggenbach (Ed.). *Perspectives on fluency* (pp. 102-127). Ann Arbor, MI: The University of Michigan Press.

Wiese, R. (1984). Language production in foreign or native languages: Same or different? In H. Dechert, D. Mohle, & M. Raupach (Eds.). *Second language productions* (pp. 11-25). Tubingen, Germany: Gunter Narr Verlag Tubingen.

Zukowski-Faust, J. (1984). Problems and strategies: An extended training program for foreign teaching assistants. In K. Bailey, F. Pialorsi, & J. Zukowski-Faust (Eds.). *Foreign teaching assistants in U.S. universities* (pp. 76-86). Washington, D.C.: National Association for Foreign Student Affairs.