Using synchronous speech to facilitate acquisition of English rhythm: A small-scale study

ELINA BANZINA[a1], LYNNE E. HEWITT[b], LAURA C. DILLEY[c]
[a]Stockholm School of Economics in Riga
[b]Bowling Green State University
[c]Michigan State University

ABSTRACT

EN While appropriate stress and rhythm is of importance for any speaker's intelligibility, such properties are critical for international teaching assistants (ITA), who deliver new and complex information to native speaker audiences. Given the limited time available for ITA instruction and the need for a time-efficient rhythm teaching method, this article reports findings of a small-scale feasibility study that tested the effectiveness of a synchronous speech component introduced into conventional rhythm instruction. Synchronous speech involves teacher and learner speaking in unison continuously, which allows the L2 learner to learn rhythm implicitly and uninterruptedly, and provides rich auditory-visual input, ample motor speech practice and real-time feedback, thereby automatizing rhythm patterns. In a 6-week-long pre-post experimental feasibility study, blind listeners evaluated pre-training and post-training recordings of ITA-produced speech. Data revealed a trend towards more improvement in L2 rhythm working with the synchronous speech technique. Results establish feasibility in both instruction and research.

Key words: SPOKEN ENGLISH, PRONUNCIATION TEACHING, ENGLISH RHYTHM, SYNCHRONOUS SPEECH, RHYTHM INSTRUCTION.

ES Si un acento y un ritmo apropiados son fundamentales para la comprensión de cada hablante, estas propiedades son indispensables para los asistentes internacionales de la enseñanza (ITA), quienes proporcionan informaciones nuevas y complejas a un público de hablantes nativos. Dado el poco tiempo del que disponen los ITA para su formación y la necesidad de un método de enseñanza eficaz, este artículo transmite los resultados de un estudio realizado a pequeña escala, en el que se ha comprobado la eficacia de la introducción del discurso síncrono como complemento del tradicional método de enseñanza del ritmo. Un discurso síncrono implica que profesores y estudiantes hablen al unísono de forma continua, lo que permite al estudiante de L2 aprender el ritmo de manera implícita e ininterrumpida gracias a un fuerte input auditivo y visual, un incremento de la práctica motora y una retroalimentación inmediata, que conducen a la automatización de los esquemas rítmicos. Para este estudio, unos auditores anónimos evaluaron las grabaciones de discursos de los ITA antes y después del entrenamiento. Los resultados muestran una notable mejora del ritmo en L2 gracias a la técnica del discurso síncrono y su factibilidad en la enseñanza y en la investigación.

Palabras clave: INGLÉS HABLADO, ENSEÑANZA Y PRONUNCIACIÓN, RITMO INGLÉS, DISCURSO SÍNCRONO, ENSEÑANZA RITMO.

IT Se accento e ritmo appropriati sono fondamentali per la comprensione di ogni parlante, tali proprietà sono indispensabili per gli assistenti internazionali all'insegnamento (ITA), che devono fornire informazioni nuove e complesse a parlanti nativi. Dato il poco tempo a disposizione per la formazione degli ITA e la conseguente necessità di un metodo d'insegnamento efficiente, questo articolo riporta le conclusioni di uno studio su piccola scala con cui si è testata l'efficacia dell'introduzione del parlato sincrono all'interno del tradizionale metodo di insegnamento del ritmo. Il parlato sincrono prevede che insegnanti e studenti parlino all'unisono e senza interruzioni, permettendo agli studenti di L2 di apprendere il ritmo in modo implicito e continuo grazie al ricco input visivo e auditivo, all'incremento della pratica motoria e al feedback immediato, consentendo di automatizzare gli schemi ritmici. Nel corso dello studio, durato sei settimane, ascoltatori esterni hanno valutato registrazioni di parlato degli ITA prima e dopo il training. I risultati mostrano un considerevole miglioramento del ritmo nella L2 grazie alla tecnica del discorso sincrono e dimostrano la sua fattibilità sia nell'insegnamento sia nella ricerca.

Parole-chiave: INGLESE PARLATO, INSEGNAMENTO PRONUNCIA, RITMO INGLESE, DISCORSO SINCRONO, INSEGNAMENTO RITMO.

---

[1] Contact: elina.banzina@sseriga.edu

1. Introduction

Appropriate stress and rhythm patterns have been shown to be critical for non-native speakers' intelligibility in numerous research studies (Anderson-Hsieh, Johnson, & Koehler, 1992; Derwing, Munro, & Wiebe, 1998; Tajima, Port, & Dalby, 1997). This is especially true in extended discourse, such as giving a presentation or teaching a class, which is overall a cognitively demanding context (Hahn, 2004). Use of natural stress and rhythm patterns is also potentially important under degraded listening conditions, such as in large lecture halls and with background noise in the classroom (Mattys, Brooks, & Cooke, 2009; Munro, 1998; Woodford, Prichard, & Jones, 1998). While this applies to any non-native speaker, the need for producing appropriate speech patterns is especially critical for international teaching assistants (ITAs), who often work with relatively large American undergraduate student audiences in an acoustically challenging environment. However, international graduate students often struggle with rhythmic performance. There is evidence in second language research that non-native patterns of rhythm and stress are pervasive in the speech of second language learners (Anderson-Hsieh et al., 1992; Flege & Bohn, 1989; MacAllister, Flege & Piske, 2002; White & Mattys, 2007). These non-native rhythmic patterns may result in seriously decreased intelligibility, or at least reduce the speed with which the message is processed by native listeners (Anderson-Hsieh & Koehler, 1988; Braun, Lemhofer, & Mani, 2011; Field, 2005; Hahn, 2004). Unfortunately, time for English rhythm training for international teaching assistants is often limited, necessitating a search for an effective approach that would lead to automaticity in a short amount of time, maximizing the amount of both input and active practice.

The synchronous speech technique, where an instructor and a student engage in reading of the same text or reciting it aloud simultaneously, was observed in our practice to yield improvement in rhythmic performance in familiar and practiced contexts. More importantly, it led to a considerable improvement also in spontaneous rhythm use in cognitively demanding contexts. The perceived benefits of the synchronous speech technique for second language speech learning are the simultaneously provided (i) rich auditory-visual input, (ii) real-time implicit feedback, (iii) ample motor speech practice, and (iv) uninterrupted flow. In fact, simultaneity is one of the key factors that allows multiple goals to be accomplished in a short amount of time, potentially leading to a long-lasting and consistent use of English rhythm and stress.

Despite the observed benefits of this method, no empirical evidence exists regarding the effectiveness of the synchronous speech technique for language rhythm acquisition. We therefore attempted to establish its applicability to rhythm instruction experimentally, in a 6-week-long small-scale experiment where synchronous speech was incorporated into an existing conventional pronunciation instruction in an oral proficiency program at a communication disorders department of a Midwestern U.S. university. The synchronous speech technique, as it was used in this study, was geared towards the needs of the particular audience—international graduate students (ITAs)—who had to prepare for giving oral presentations and lecturing in front of undergraduate students. In our 6-week experiment, students in the experimental group were not taught exclusively with the synchronous speech technique but rather a synchronous speech element was *added* to the conventional instruction so that students would receive explicit teaching as well. It is generally agreed in second language research that adult learners benefit both from implicit and explicit instruction (Ellis, 2005, 2007; Hulstijn, 2005). While synchronous speech allows rhythm to be learned implicitly, having metalinguistic knowledge of the English language rhythm is beneficial in order to notice rhythmical differences more easily and be able to respond to corrective feedback more effectively. Moreover, the experimental group was an existing class and we did not want to employ a new and previously untested instructional technique as our sole approach. We therefore added a synchronous speech component to the traditional explicit rhythm instruction. This allowed for a quasi-experimental design, examining whether the addition of this component resulted in better performance in that group.

The current study sought primarily to answer whether it is experimentally feasible to test the effectiveness of the synchronous speech technique applied in pronunciation teaching and what measures might be appropriate to establish its effectiveness. The other main goal of this study was to determine whether the use of the synchronous speech technique can lead to improved rhythmic performance in a second language in spontaneous production. While the focus of this study was on rhythm teaching and learning, we also aimed to identify other areas of oral performance that might benefit from synchronous speech practice. Although the current small-scale study cannot establish definitive effectiveness of the technique, we sought to garner preliminary data supporting its potential effectiveness, both in enhancing performance in the synchronous contexts, as well as transferring improved performance to spontaneous

speech. Thus, the present investigation is a feasibility study, working out a protocol to test the method and to establish whether preliminary results would justify further testing.

2. Literature Review

Languages differ from one another in terms of prosody and rhythm. The rhythm of the English language features alternations of more prominent and less prominent syllables, the more prominent ones demonstrating greater intensity, fundamental frequency, and duration than the less prominent ones, which often get reduced (Fry, 1955; Ladefoged, 2005; Lehiste, 1970). Production studies have consistently shown that L2 learners' production of these suprasegmental features in English deviates from native speakers' performance. This most often is a reflection of non-native speakers' experience with patterns and features that are employed and viewed as critical in their first language phonological system and subsequently, transferred to learners' L2, where they are continued to be used in L1-specific ways. For example, prior research has shown that L2 learners tend to apply phonological knowledge of their native language when perceiving and producing stressed and unstressed syllables in English (Chen, Robb, Gilbert, & Lerman, 2001; Lee, Guion, & Harada, 2006; Yu & Andruski, 2010), duration cues (Lehiste & Fox, 1992) and vowel reduction (Braun, Lemhöfer, & Cutler, 2008; Flege & Bohn, 1989; Zhang, Nissen, & Francis, 2008).

The use of such non-native patterns in L2 speech can have detrimental effects on speech intelligibility, all the more so in instructional speech that delivers novel and cognitively demanding content. L2-accented instructors are the reality in many U.S. universities that are increasingly reliant on International Teaching Assistants for filling the instructional needs of U.S. undergraduate students (Chiang, 2009). The quality of instructors' oral delivery should ensure access to the academic content maximally quickly and effectively, yet high numbers of undergraduate students experience difficulties understanding or connecting with their instructors' speech (Pickering, 2001; Rao, 1995). According to Stevens (1989), students cite the non-native use of suprasegmentals as the main reason for the failure to decode their instructors' messages. Suprasegmental accuracy has been long viewed as more critical for comprehension than segmental accuracy in L2 pronunciation literature (Anderson-Hsieh et al., 1992; Derwing, Munro, & Wiebe, 1998; Morley, 1991). Some of the most common problem areas addressed in ITA research literature have been improper use of stress, insufficient vowel reduction, inadequate intonation patterns, pausing, and speech rate (Gorsuch, 2013; Hahn, 2004; Kang, 2008; 2010; Pickering, 2001), which have interfered with undergraduate students' performance at some level. The time for speech training in the ITA training curriculum is unfortunately limited and often needs to be shared with other aspects of ITA training such as culture and pedagogy (Anderson-Hsieh, 1990). This calls for highly effective pronunciation teaching techniques that yield immediate improvements in the suprasegmental realm. Some of the teaching approaches proposed for classroom teaching of suprasegmentals have been, for example, a cognitive approach with a metacognitive analysis of the suprasegmental system and subsequent self-study using field-specific materials (Anderson-Hsieh, 1990) and a drama-based approach (Stevens, 1989). Often, the focus has been entirely on self-study after pronunciation instruction, such as an input-based approach in the form of repeated reading along with an audio recording (Gorsuch, 2011), or covert practice with self-monitoring and self-correction (Dickerson, 2000). However, in order to maximize class time, the search for most effective teaching techniques continues.

The current study was designed to provide empirical evidence for the effectiveness of the synchronous speech technique, which has been anecdotally observed to yield positive results in L2 rhythm acquisition. The exploration of this technique was motivated by the empirical accounts of synchronous speech being successfully used to achieve the desired speech effect in the field of communication disorders, where it is used in stuttering therapy for speech rate, prosody control, and fluency development (Freeman & Armson, 1998; Ham, 1988; Kalinowski & Saltuklaroglu, 2003; Kiefte & Armson, 2008). While L2 speech acquisition is an entirely different discipline and of a non-clinical nature, the shared speech intelligibility goal justifies further explorations of professional practices used in the speech disorder discipline as part of an interdisciplinary quest for most effective instructional tools.

In the literature, the simultaneous speech production by two speakers has been referred to in various ways: synchronous speech (Cummins, 2003), choral speech or chorus repetitions (Freeman & Armson, 1998; Kjellin, 1999), and unison speech (Ham, 1988). It is important to note two things: first, the technique described here does not involve rhythmical, highly structured chanting, such as reciting poems and rhymes, but rather refers to any other less structured text that one would encounter in their daily lives. Second, synchronous speech, as its name suggests, really denotes speech that is created by two people simultaneously, without a significant lag in production. This is different from speech shadowing, a related yet

different term, in which one speaker attempts to reproduce the other speaker's speech after it is uttered, without a text in front of them, and with no familiarity with the upcoming utterance (Cummins, 2003). While Cummins (2003) points out that there are lags also in a synchronous speech condition, they are much shorter than the shortest lags in shadowed speech. Kalinowski and Saltuklaroglu (2003) describe the timing of synchronized speech as "one signal imitating the other in a loosely defined temporal window" (p. 342).

Kjellin (1999) has provided an experience-based account of introducing synchronous speech into second language pronunciation instruction, where relatively short phrases were used for teaching prosody. He notes that choral repetition is not passive audio-lingual drilling, but rather efficient active work both mentally and physically automatizing a skill, the instant results of which generates a high level of motivation in the student, so repetitions never get boring. Chela-Flores (1994) gives indirect support to the synchronous speech technique in second language rhythm acquisition by pointing out the need for the student to develop a feel for the continuous flow of rhythm, rather than working on rhythm on a word-by-word basis that does not let the learner perceive it as a whole and form a general habit of its use. In such cases, rhythm use is limited only to the learned words and phrases and does not generalize to out-of-class spontaneous speech situations. Similarly, she emphasizes the difference between those mechanisms that lead to conscious knowledge of language rules but do not normally develop oral fluency, and those mechanisms that develop automatic language competency, claiming that in the traditional rhythm instruction more attention is usually given to the former.

"Conventional" rhythm teaching typically combines both explicit instruction and meaningful and communicative practice to gradually move from the declarative stage of skill acquisition to the procedural stage and finally to the automatic stage (DeKeyser, 2007), which is a relatively slow process. Synchronous speech could hypothetically accelerate rhythmic automaticity by focusing on the final automatization phase, where rhythmic patterns would be internalized through their smooth and fluent use. The intensive nature of synchronous speech provides frequent perceptual exposures to the feature learned and a considerable amount of motoric practice. Thus the synchronous speech technique could minimize reliance on the metalinguistic rules of rhythm production and the procedural stage of rhythm learning, instead allowing the learners to fully immerse themselves into "assisted" native-like rhythm use.

A neurological perspective on synchronous speech could explain why instant imitation of native speaker's speech gestures leads to automaticity and could potentially be a more effective tool for rhythm acquisition than more analytical approaches. Kalinowski and Saltuklaroglu (2003) articulate the position that all humans possess "mirror neurons" that create an immediate link between an observation and an action (or perception and production) and allow humans to imitate complex motor sequences. While the status of mirror neurons remains unclear and somewhat controversial, recent reviews indicate that there is strong evidence to support their existence and role in motor imitation and hence motor learning (Molenberghs, Cunnington, & Mattingley, 2012). Thus synchronous speech could be conceptualized as a simple input-output reflex mediated by mirror neurons, during which "the only cognitive tasks required are perceiving and attending to the external signal during speech production" (Kalinowski & Saltuklaroglu, 2003, p.344). Regardless of whether such a strong claim can be substantiated, the mere presence of a mirror neuron system can be seen as evidence for the neurological importance of imitation in human motor learning.

In sum, the strength of the synchronous speech technique lies in the fact that rhythm is learned (i) psycho-motorically and implicitly, (ii) uninterruptedly and holistically, (iii) intensively and time-efficiently, while (vi) maintaining consistency across speech. The flow of rhythm develops in time; interrupting the flow by the instructor and the student taking turns would be detrimental to rhythm learning. Therefore, the continuity and fluency of practice gives students an opportunity to use the desired rhythm in a native-like and fluent manner, forming a strong psycho-motoric habit without much conscious awareness. Additionally, this lets students hear themselves speak according to native prosodic norms, which might be a powerful motivator. Moreover, temporal features of rhythm are hard to maintain without a real-time metronome that is instructor's speech, and could be lost if not paced and aligned by synchronous speaking. The outside real-time meter helps students set their "inner clock". Finally, automatizing a skill is the final and most time-consuming stage of skill acquisition, yet this final stage of skill learning is often cut short due to time constraints. The synchronous speech technique puts emphasis on the last step — the skill automatization process — from the very first day, and through fluent and intensive use could arguably turn the new feature into a habit faster.

3. Methods

3.1. Research design

The small-scale feasibility study used a pre-test post-test control group quasi-experimental design, where the control group (n = 3) worked with the regular rhythm instruction, and the experimental group (n = 3) worked with a synchronous speech component added to the regular instruction. The research participants were audio-recorded at the beginning and end of the 6 weeks of rhythm instruction period to obtain (i) initial baseline speech samples and (ii) final post-instruction speech samples. The last ones assessed the potential transfer to novel contexts/spontaneous speech as students were giving short unscripted presentations.

The baseline and final recordings were further subjected to perceptual measures to compare progress in rhythmic performance in each group. Recordings from all six participants were first low-pass filtered with speech analysis software to retain only suprasegmental information, in order to eliminate any potential bias from context factors, and further perceptually rated by three certified, licensed speech-language pathologists, all naïve to the study, on (i) rhythm pattern acceptability, (ii) speech rate acceptability, and (iii) overall speech acceptability for instructional purposes.

3.2. Participants

The study was conducted within a 6-week-long course of oral proficiency for international graduate teaching assistants (ITAs) in a department of communication sciences and disorders at a university in Midwestern United States. In this course, each international student received individual instruction by a speech-language pathology graduate student, supervised by a licensed and certified speech-language pathologist.

Two groups of participants were involved in the study: first, six international students (one female, five male), and second, six native-English speaking instructors (all female) who all volunteered to participate in the experiment. The international students had all been screened for English language oral proficiency by ESL experts during the students' first days on campus. If their speech patterns did not meet the established fluency criteria, they were further referred to the department of communication sciences and disorders for placement in their oral proficiency program. The supervising SLP along with the participating graduate students administered initial oral proficiency screenings to each referred student and identified individual pronunciation goals. Finally, those international students who were identified as having non-native prosodic patterns were further offered participation in the current study.

The six international participants who volunteered were randomly assigned to control and experimental groups, controlling only for the native language factor: the complexity of certain languages could affect the speed with which rhythm is acquired; therefore, participants' native languages were balanced across both groups. Thus, there were two speakers of Chinese in each group, while the remaining two participants — one speaker of Nepali and one speaker of Sinhala — were randomly assigned to the control and the experimental groups, respectively.

Table 1

Participants' native language and field of study.

+	NATIVE LANGUAGE	FIELD OF STUDY
Control 1	Chinese	Chemistry
Control 2	Chinese	Chemistry
Control 3	Nepali	Geology
Experimental 1	Chinese	Biology
Experimental 2	Chinese	Math & Statistics
Experimental 3	Sinhala	Chemistry

3.3. Procedure

Speech instructors worked with the ITAs on English rhythmic patterns for a total of six consecutive weeks during the semester, where approximately up to 30 minutes out of the regular 50-minute class each week were devoted to teaching rhythm and stress. Although each session was conducted individually, maximum measures were taken to make sure that all instructors performed the same way through intensive training and constant monitoring. Before meeting with their students, all native-speaking instructors completed weekly joint training sessions where they practiced synchronous speech with each other and shared their experiences. Instructors were provided with detailed guidelines and initial texts to use for the rhythm teaching part of each session.

Each instructor met with his or her assigned ITA in a quiet therapy room equipped with a video camera. One of the researchers constantly monitored the instruction process via camera on a computer in another room, and gave instant feedback and made the necessary corrections right after the session to ensure fidelity of treatment.

Both control and experimental groups spent an equal amount of time per session on rhythm instruction. Within each session, all students went through controlled, guided and free speech activities (see Celce-Murcia, Brinton & Goodwin, 2006); however, the proportions of these in each group were different. The experimental group received proportionally fewer guided and free speech activities, which were substituted with synchronous speech activities instead.

The explicit instruction on rhythm that both groups received involved work on word stress (stress realization: duration, pitch, loudness, as well as vowel reduction), primary phrase stress, content and function words, thought groups, stress in compound nouns, and prominence in discourse context (stressing new information versus old, emphatic stress, contrastive stress). Each session began with word- and short phrase-level controlled activities with ear training (Celce Murcia et al., 2006) and modeling by the instructor and imitation by the student. The following guided activities involved more field-specific content and longer utterances; still, the attention of the student was focused on rhythm production. Feedback at this stage was explicit and immediate. Finally, the free speech and communicative activities were aimed at developing automaticity and use of the newly acquired skills in spontaneous and also cognitively demanding speech, which is an everyday reality for teaching assistants. These activities were aimed at fostering presentation-type of speech and were made maximally meaningful to the ITAs; for example, they involved teaching some basic concepts in their field, describing research conducted in their department, giving a motivational speech for prospective students in their department, etc. Feedback at this stage could be instant or delayed.

Both experimental and control groups received the abovementioned instruction, but only the experimental group received instruction using the synchronous speech technique in place of some guided and free speech activities. The experimental group followed specific detailed guidelines for work with synchronous speech. The materials used working with the experimental technique were field-specific texts and presentation materials that were either selected for the student or, later in the semester, provided by students themselves (e.g., descriptions of research interests, journal abstracts, upcoming presentations). Using contexts relevant to students' academic needs was an approach selected to enhance student motivation. The texts used consisted of whole paragraphs, and due to their length could be repeated several times only. Nevertheless, the length of the material also meant that the flow of rhythm could be internalized more continuously and intensively.

Initially, each synchronous speech activity began with the instructor modeling for the student and the student silently lip-reading along to familiarize him or herself with the text. Next, the student joined in, and both student and instructor spoke in unison. Depending on the length of the text, it was repeated three to five times, with the instructor prompting the student to gradually establish eye contact with the "audience", i.e., the instructor. This was done in order to minimize reliance on the script and to create an impression of a real presentation. Looking up frequently at the imaginary audience also helped maintain thought groups and appropriate pausing in the running speech. Often, the student and the instructor would stand in the opposite corners of the therapy room to create distance between the speaker and the audience, creating the effect of a real presentation where appropriate eye contact and appropriate volume needs to be maintained; this also tended to increase the prominence of stressed syllables in speech. Additionally, to minimize the "crutch" that the instructor's voice might have turned into during the synchronous speech activity, the instructor would significantly lower the volume of his/her voice at times and let the student take the lead so students could hear themselves produce speech in a near-native manner by themselves. Afterwards, the instructor would tune back in and continue as usual.

It was important for the instructors to attend to several things while working with this technique: first, they had to project their voices loudly enough that the student could speak in full voice and still hear the instructor' voice, which was not always easy. Second, the instructors were trained to use expressive body language (e.g., raising eyebrows, looking up, leaning forward) to amplify rhythmic patterns. Body language is of benefit in such instruction since it allows speech synchronization through peripheral vision while reading a text (Cummins, 2003), alerts the student that a new thought group is coming, and, overall, serves as a reminder of the physiological aspect of stress production. In this way, body language acts as a real-time feedback mechanism, since it can convey the right timing and prominence without using voice or any other signal. Overall, corrective feedback was implicit, delivered real-time via amplified auditory and visual cues. For example, if the student rushed ahead, the instructor could slow down her/his speech even more or prolong the stressed syllables if they were of insufficient duration in the student's production.

3.4. Data Collection

To establish the efficacy of the synchronous speech technique, pre-test and post-test recordings were obtained from each group. Recordings were made with Panasonic RR-US470 digital audio recorders and SHURE C606 professional microphones attached to them, which were distributed to all instructors. Overall, students were recorded six or more times during the semester, but only the first and the last recordings were used in this study as we were interested in the maximum result each technique could achieve in the limited amount of time allocated for rhythm instruction. Students were unaware of which recordings would be used for research purposes, and recordings were regularly used in class for work with segmentals, so students were relatively comfortable with being recorded.

For baseline recordings, all participants read a scripted text (see Appendix A). They were asked to first familiarize themselves with the text and read it silently, and then read it and say each line aloud, looking up, as if presenting it to their class. For the final recordings that assessed retention and transfer to spontaneous production, students were recorded as they produced a novel and spontaneous speech that they had a few minutes to prepare. Prior to recordings, students were told to remember the segmental and suprasegmental features they had worked on during the semester and use them accurately in their final speech; the very final session was devoted entirely to assessment and involved no rhythm instruction. The speech had to be on some concept from their field they could explain and then students were asked to "present" it while they were being recorded. The strong focus on the content made attending to speech patterns cognitively demanding, as students had to construct a novel text in their minds and simultaneously maintain appropriate rhythm. Since most of the cognitive resources were spent on the challenging context, leaving little for attending to the rhythmic performance, the ability to maintain rhythm in the spontaneous and content-wise difficult speech was taken as an indicator that a certain degree of automaticity had been achieved.

3.5. Data Analysis

Data were analyzed using descriptive statistics to assess the outcomes of the synchronous speech component in the experimental group and establish research feasibility of this method for future studies. The small sample size of this study did not allow the use of inferential statistical methods or the ability to generalize results to larger populations. Selected recorded speech samples from all six participants were randomized and presented to three naïve speech-language pathologists who assigned impressionistic ratings to the speech in terms of (i) rhythm pattern accuracy, (ii) speech rate acceptability, and (iii) overall speech acceptability as instructional speech. Regarding the latter, we hypothesized that there might potentially be additional factors besides the first two that could influence the perceived acceptability of speech.

All speech samples were approximately 30 seconds long, and contained not less than five sentences. Such sample length was deemed to be sufficient to appropriately reflect the target patterns and decrease the burden on raters. Since the baseline recordings were all done on the same text for all speakers, the 30-second portion was randomly spliced out of various time points so that the content would not repeat itself and raters would not develop familiarity effects with speech that would make them rate samples later in time more favorably. The final recordings, on the other hand, were recorded on spontaneous speech that was different for each individual speaker; therefore, the 30-second speech samples were spliced out from the middle of the recording for all participants.

To ensure that raters only focused on the suprasegmental aspect of speech and interference of segmental information was eliminated, all speech samples were low-pass filtered at 1000 Hz using Praat

speech analysis software (Boersma & Weenik, 2011). Low-pass filtering passes lower frequencies and rejects higher frequencies, making segmental information relatively unavailable to the listener as words are almost impossible to make out. Suprasegmental information, on the other hand, was all preserved and thus the only quality information available to raters.

Speech samples were presented to raters on a computer via headphones. A 100 mm long visual rating scale was used, where raters selected a point on the continuous blank scale, marked as, for example, "non-native rhythm" at one end, and "native-rhythm" at the other end (see Appendix B). Raters were told to make use of the whole scale. Care was taken to make sure that all raters defined the notion of "rhythm" identically and were able to apply it to the ratings; therefore, they received a short introduction to the rhythmical system of the English language, followed by a short training session on random and study-unrelated low-pass-filtered native and non-native speech samples.

Raters could listen to each speech sample twice and they then marked the scale according to the perceived performance in English. The inter-rater reliability was high, r (34) =.71, p < .01. Intra-rater reliability on the same speech samples that raters had to rate twice was even higher, r (16) = .93, p < .01). Each individual ITA's progress during the semester was expressed as a difference in scaled units (0 - 100 on the visual rating scale) between the baseline recording and the final recording (i.e., final perceptual rating minus initial perceptual rating), averaged across all three raters.

4. Results

The two groups' performances were rated on a scale of 1 to 100 for rhythm accuracy in spontaneous oral production, as well as acceptability of speech rate and overall speech patterns in spontaneous production. An individual differences analysis revealed that one of the participants in the Experimental group exhibited a profile that matched neither the control nor experimental group: her performance at the end of the semester was well below the average in any of the groups; therefore, this participant was excluded from further analyses.

Figure 1 presents results of the perceptual improvement in rhythmic pattern use in spontaneous speech, which was the focus of this study. The amount of improvement in the control group (n = 3) was an average of 18.4 units, while in the experimental group (n = 2) the improvement was an average of 30.7 units. This indicates that, perceptually, the experimental group experienced greater improvement in second language rhythm use in spontaneous production than the control group, potentially due to the addition of the synchronous speech technique.
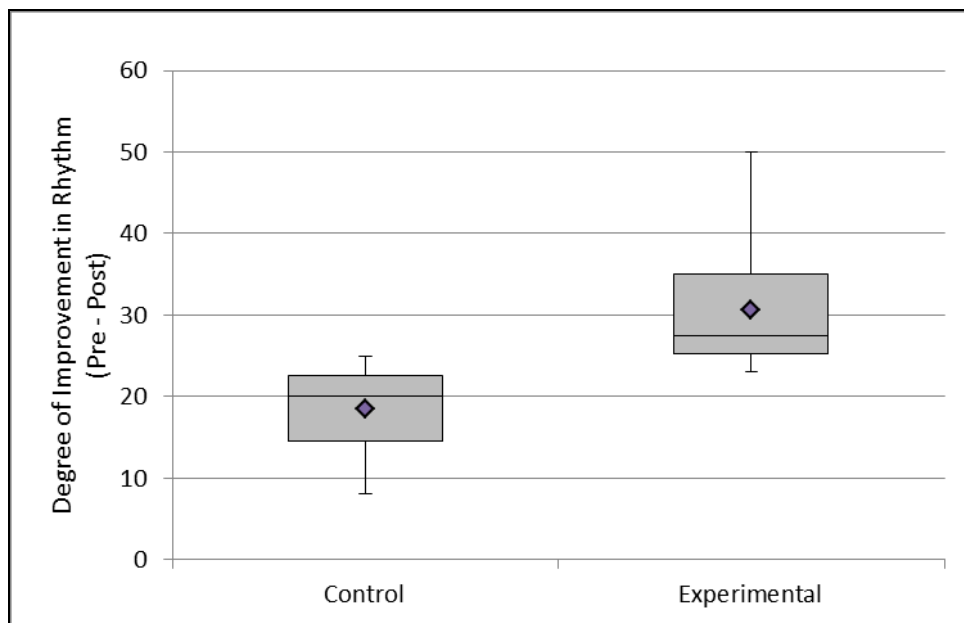

Figure 1. Difference in speech rhythm acceptability ratings pre- and post-instruction on a scale from 0 to 100, indicating perceived improvement in speech rhythm in control and experimental groups.

Figure 2 illustrates perceptual ratings for improvement in the overall speech patterns. A slight trend toward greater overall speech improvement in the experimental group (*M* = 33.6), compared to the control group (*M* = 19.3), can be observed.
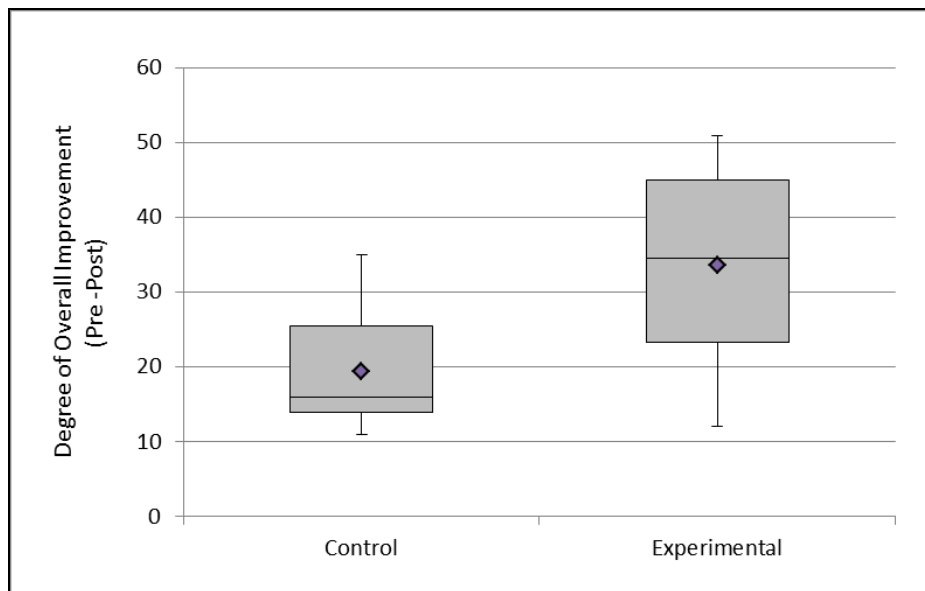

Figure 2. Difference in overall speech pattern acceptability ratings pre- and post-instruction on a scale from 0 to 100, indicating perceived improvement in overall speech patterns in control and experimental groups.

Finally, while both groups were rated as having made comparable progress in improving speech rate, improvement was more pronounced in the experimental group (*M* = 37.5), compared to the control group (*M* = 20.5), again showing a trend towards greater benefits of adding the synchronous speech technique.
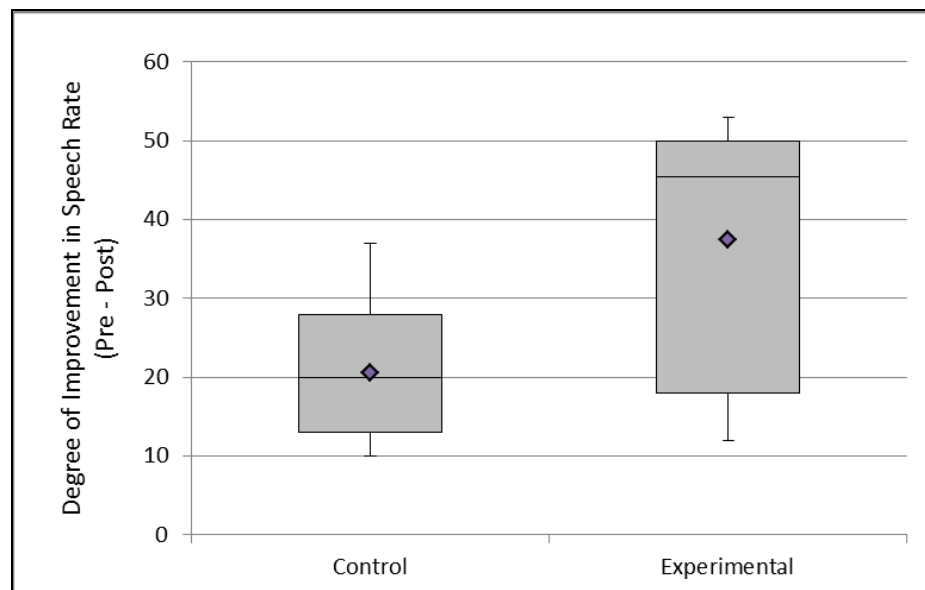

Figure 3. Difference in speech rate acceptability ratings pre- and post-instruction on a scale from 0 to 100.

5. Discussion and pedagogical implications

The number of international graduate students teaching at U.S. universities is large and growing. Interactions between ITAs and American undergraduates have often resulted in communication failures, which is hardly surprising given that the amount of time allotted to ITA preparation, speech training in particular, typically is short. Various approaches for teaching English speech patterns have been proposed and tested empirically, but the synchronous speech technique so far has not been addressed in research literature. This study attempted, first, to establish the feasibility of testing the effectiveness of the synchronous speech technique with experimental methods and develop a protocol to test the method, and second, to provide preliminary experimental evidence for the effectiveness of the synchronous speech technique in L2 rhythm teaching. The hypothesized benefits of using synchronous speech in pronunciation instruction were the simultaneously available rich input and practice, as well as implicit, fluent and continuous rhythm learning that could potentially allow for automatizing rhythm production faster than other rhythm teaching methods and techniques.

Out of the six volunteers who agreed to participate in our small-scale study conducted within an existing pronunciation course, one participant was eventually eliminated due to exhibiting performance that was well below the average for both the experimental and control groups, apparently failing to benefit from any type of instruction. There could be a range of possible explanations here, including the fact that this course was mandatory for all six potential teaching assistants and certain motivation-related issues might have played a role.

The study findings with the five remaining participants partially confirmed our casual observations and intuitions regarding the effectiveness of the synchronous speech technique: observable differences between the experimental and control groups were found, such that there was a trend toward more pronounced improvement in speech rhythm acquisition when the synchronous speech technique was added to instruction. Considering the small sample size (six participants total, out of which one was eliminated as an outlier) and the limited time for instruction (a total of six weeks, not more than half an hour per week), it can be considered an encouraging preliminary finding. An added synchronous speech component thus has shown potential to accelerate learning of second language rhythm patterns and aid their transfer to new and untrained contexts. However, most surprising and encouraging might be the fact that these rhythmic improvements were shown in spontaneous speech production, which is the ultimate goal of pronunciation training. What is more, improved rhythmic performance was demonstrated in cognitively demanding contexts, where most of the cognitive resources were devoted to the content of the speech. These results indicate that automaticity in second language rhythm use may be achieved faster with the addition of the synchronous speech technique than with conventional rhythm teaching techniques alone.

Initial screenings indicated that speech rate normalization, specifically its reduction, was among the many semester goals set for these students. Although trending toward greater improvement, the application of the synchronous speech technique did not appear to help with speech rate normalization considerably more than the "conventional" rhythm teaching techniques. Neither baseline nor final ratings as assigned by the three raters reached the maximum score; therefore, the possibility of a ceiling effect was eliminated. One possible explanation for the clear lack of difference between groups could be that the time spent on the presentation type of speech for both groups, irrespective of the type of rhythmic training received, led to increased awareness of the needs of the audience and consequently speech rate normalization.

Similarly, the experimental group was not perceived by raters as having convincingly more improvement in the overall speech patterns (intonation, pausing and other factors included) than the control group, although a trend toward more improvement was again observed. Possibly, the amount of time devoted to working with the synchronous speech was insufficient to reveal its full potential in this realm. The small sample size and unequal number of participants limit any conclusions in this regard.

From our preliminary results, we can tentatively suggest that one aspect of second language speech learning where synchronous speech could have superiority over traditional techniques is rhythm pattern acquisition, which was the focus of this study. More importantly, the findings for the experimental group indicate that 30 minutes of instruction over six weeks might be sufficient to see the benefits of the synchronous speech technique for automating second language rhythm patterns. The presence of the outlier who decreased her rhythmic accuracy and speech acceptability following six weeks of instruction is a particular complicating factor in interpreting the present results. Thus a larger-scale study with a greater sample size and a longer instruction period is needed.

Overall, our results supported the feasibility of testing the technique experimentally by using the selected measures of improvement and the efficacy of this method for rhythm training in general. The use of the low-pass filtered speech and impressionistic ratings as a measure proved to be sensitive enough to detect differences in the rhythmic improvement after a limited number of training sessions. The fact that the experiment was conducted within an existing class dictated that synchronous speech in the experimental group not be used in isolation but in addition to more explicit rhythm teaching instruction to minimize risks for the students; as a result, the ability for synchronous speech to achieve the current result independently is yet unknown. Theoretically, the explicit component of rhythm teaching that is needed to help students notice second language features could be substituted with real-time amplification of rhythmic cues during synchronous speech, thus making rhythmic acquisition possible without referring to metalinguistic means of instruction.

This study also demonstrated that synchronizing speech for native and non-native speakers in an instructional setting was feasible, dispelling our initial concerns. Although Cummins (2002, 2003) observed that when two people recite a common text in unison, they rapidly attune their speech to each other and are able to successfully synchronize their speaking with relative ease, these observations were limited to two native speakers who share the same language and expectations regarding its prosody. This study proved that such observations can be extended also to native—non-native speaker dyads; however, it does not allow for generalizations regarding the effectiveness and implementation of the technique in larger groups, which would require various adjustments and more creativity on the part of the instructor.

Although the population that was targeted with the synchronous speech technique in this study was international teaching assistants, its application is not limited to ITAs or academic settings but can be successfully used in a variety of second language teaching contexts where a more holistic approach to rhythm teaching is preferred and time for speech instruction is limited.

References

Anderson-Hsieh, Janet (1990). Teaching suprasegmentals to international teaching assistants using field-specific materials. *English for Specific Purposes, 9*, 195-214.

Anderson-Hsieh, Janet, Johnson, Robert, & Koehler, Kenneth (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning, 42*(4), 529-555.

Anderson-Hsieh, Janet, & Koehler, Kenneth (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning, 38* (4), 561–613.

Braun, Bettina, Lemhöfer, Kristin, & Cutler, Anne (2008). English word stress as produced by English and Dutch speakers: The role of segmental and suprasegmental differences. Proceedings from INTERSPEECH 2008: *9th Annual Conference of the International Speech Communication Association*. Brisbane, Australia, p.1953.

Braun, Bettina, Lemhöfer, Kristin, & Mani, Nivedita (2011). Perceiving unstressed vowels in foreign-accented English. *Journal of the Acoustical Society of America, 129* (1), 376–387.

Boersma, Paul, & Weenink, David (2011). Praat: doing phonetics by computer [Computer program]. Version 5.2.16, retrieved 20 February 2011 from http://www.praat.org.

Celce-Murcia, Marianne, Brinton, Donna, & Goodwin, Janet (1996). *Teaching pronunciation: Reference for teachers of English to speakers of other languages*. Cambridge, United Kingdom: Cambridge University Press.

Chela-Flores, Bertha (1994). On the acquisition of English rhythm: Theoretical and practical issues. *International Review of Applied Linguistics*, *32*(2), 232-242.

Chen, Yang, Robb, Michael P., Gilbert, Harvey R., & Lerman, Jay W. (2001). A study of sentence stress production in Mandarin speakers of American English. *Journal of the Acoustical Society of America, 109*(4), 1681–1690.

Chiang, Shiao-Yun (2009). Dealing with communication problems in the instructional interactions between international teaching assistants and American college students. *Language and Education, 23*(5), 461-478.

Cummins, Fred (2002). On synchronous speech. *Acoustic Research Letters Online, 3*(1), 7-11.

Cummins, Fred (2003). Practice and performance in speech produced synchronously. *Journal of Phonetics, 31*, 139–148.

DeKeyser, Robert (2007). Skill acquisition theory. In Jessica Williams and Bill Van Patten (Eds.), *Theories in second language acquisition: An introduction*. Mahwah, NJ: Erlbaum, 97-113.

Derwing,Tracey, Munro, Murray, & Wiebe, Grace (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*(3), 393–410.

Dickerson, Wayne (2000). *Covert rehearsal as a bridge to accurate fluency*. Presented at the 34th Annual Meeting of the Teachers of English to Speakers of Other Languages, Vancouver, BC.

Ellis, Nick C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, *27*, 305 – 352.

Ellis, Nick C. (2007). Implicit and explicit knowledge about language. In Jasone Cenoz & Nancy H. Hornberger (Eds.), *Encyclopedia of language and education* (Vol. 6, pp. 119–132). New York City, New York: Springer.

Field, John (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*, 399-423.

Flege, James E., & Bohn, Ocke-Schwen (1989). An instrumental study of vowel reduction and stress placement in Spanish-accented English. *Studies in Second Language Acquisition, 11,* 35-62.

Freeman, Kathryn, & Armson, Joy (1998). Extent and stability of stuttering reduction during choral reading. *Journal of Speech-Language Pathology and Audiology, 22*, 188–202.

Fry, Dennis B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America, 27*, 765–768.

Gorsuch, Greta J. (2011). Improving speaking fluency for international teaching assistants by increasing input. *TESL-EJ, 14*(4), 1-25.

Gorsuch, Greta J. (2013). Helping international teaching assistants acquire discourse intonation: Explicit and implicit L2 knowledge. *Journal of Teaching English for Specific and Academic Purposes, 1*(2), 67-92.

Hahn, Laura (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*(2), 201-223.

Ham, Richard E. (1988). Unison speech and rate control therapy. *Journal of Fluency Disorders, 13*(2), 115-126.

Hulstijn, Jan H. (2005). Theoretical and empirical issues in the study of implicit and explicit second language learning. *Studies in Second Language Acquisition*, *27*, 129-140.

Kalinowski, Joseph, & Saltuklaroglu, Tim (2003). Choral speech: the amelioration of stuttering via imitation and the mirror neuronal system. *Neuroscience and Biobehavioral Reviews, 27*(4), 339-347.

Kang, Okim (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38, 301-315.

Kang, Okim (2008). The effect of rater background characteristics on the rating of International Teaching Assistants speaking proficiency. *Spaan Fellow Working Papers, 6*, 181-205.

Kiefte, Michael, & Armson, Joy (2008). Dissecting choral speech: Properties of the accompanist critical to stuttering reduction. *Journal of Communication Disorders, 41*, 33–48.

Kjellin, Olle (1999). Accent Addition: Prosody and Perception Facilitates Second Language Learning. In Osamu Fujimura, Brian D. Joseph, & Bohumil Palek (Eds.), *Proceedings of LP'98 at Ohio State University, Columbus, Ohio, September 1998*. Prague, Czech Republic: The Karolinum Press.

Ladefoged, Peter (2005). *A course in phonetics* (5th ed.). Boston, Massachusetts: Heinle.

Lee, Borim, Guion, Susan G., & Harada, Tetsuo (2006). Acoustic analysis of the production of unstressed English vowels by early and late Korean and Japanese bilinguals. *Studies in Second Language Acquisition, 28*, 487–513.

Lehiste, Ilse (1970). *Suprasegmentals*. Cambridge, Massachussets: MIT Press.

Lehiste, Ilse, & Fox, Robert A. (1992). Perception of prominence by Estonian and English listeners. *Language and Speech, 35*(4), 419–434.

Mattys, Sven L., Brooks, Joanna, & Cooke, Martin (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology, 59*, 203-243.

McAllister, Robert, Flege, James E., & Piske, Thorsten (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics, 30*, 229–258.

Molenberghs, Pascal, Cunnington, Ross, & Mattingley, Jason B. (2012). Brain regions with mirror properties: A meta-analysis of 125 human fMRI studies. *Neuroscience & Biobehavioral Reviews*, *36*(1), 341-349.

Morley, Joan (1991).The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25* (3), 481-520.

Munro, Murray J. (1998). The effects of noise on the intelligibility of foreign-accented speech. *Studies in Second Language Acquisition, 20*, 139–154.

Pickering, Lucy (2001). The role of tone choice in improving ITA communication in the classroom. *TESOL Quarterly, 35*(2), 233-255.

Rao, Nagesh (1995). *The "oh no!" syndrome: A language expectation model of undergraduate negative reactions toward foreign teaching assistants*. Paper presented at the 79th Annual Meeting of the International Communication Association, Albuquerque, New Mexico.

Stevens, Scott G. (1989). A dramatic approach to improving the intelligibility of ITAs. *English for Specific Purposes*, *8*, 181–194.

Tajima, Keiichi, Port, Robert, & Dalby, Jonathan (1997). Effects of temporal correction on intelligibility of foreign accented English. *Journal of Phonetics, 25,* 1–24.

White, Laurence, & Mattys, Sven L. (2007) Calibrating rhythm: First language and second language studies. *Journal of Phonetics, 35*(4), 501–522.

Woodford, Charles, Prichard, Cheryl, & Jones, Robin (1998). Listening conditions in higher education classrooms: One method of improving them. *Education, 119*(1), 129-134.

Yu, Vickie, & Andruski, Jean (2010). A cross-language study of perception of lexical stress in English. *Journal of Psycholinguistic Research, 39*(4), 323–344.

Zhang, Yanhong, Nissen, Shawn L., & Francis, Alexander L. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America 123*(6), 4498–4513.

Appendix A

Scripted text for baseline recordings

"When world-class people are invited to campus, and they are important in your discipline, be sure to attend, if at all possible. Attend their lecture even if they are from a different discipline than yours. Such people come to give seminars and to spend a day or two with faculty and students. Often, they become events to remember. You might want to attend just to see who those people are. Go even if you are not interested in the topic. What you will hear, if you listen carefully, is how they think about the world, how they approach problems, and how they work on them. They might change your perspectives about your own areas of teaching and research.

If you want to survive in a research-oriented institution, publish as much as possible. Write most of your articles for refereed journals. Papers presented at meetings get you funds to travel. However, even if refereed, these papers don't really count for tenure, salary raise, or promotion.

Your first papers might be rejected. But remember – every paper has a market, unless they are real stinkers. If the work is sound, it will get published. Submit your papers first to the best journals. Work your way down the list if a paper is rejected. In general, publish early and often. Begin writing for publication while still in graduate school. Data show that people who publish while still in graduate school usually continue to publish at a fast rate after they graduate – unlike those who did not publish while still a student."

(Adapted from Gray, P., & Drew, D. E. (2008). *What they didn't teach you in graduate school: 199 helpful hints for success in your academic career.* Stylus Publishing: Sterling, Virginia.)

Appendix B

Visual Rating Scale

RATER'S FORM

Place an X on each line to assess each speaker's rhythm, speech rate and overall speech patterns:

Participant: _____

RHYTHM

Non-native ———————————————————————— Native

SPEECH RATE

Not Acceptable ———————————————————————— Acceptable

OVERALL SPEECH PATTERNS

Not Acceptable ———————————————————————— Acceptable

Elina Banzina, Department of Languages, Stockholm School of Economics in Riga – Riga, Latvia

elina.banzina@sseriga.edu

EN	**Elina Banzina,** Ph.D., is an Assistant Professor at the Department of Languages of Stockholm School of Economics in Riga. Her research interests involve English phonetics, perception and production of foreign-accented speech, intelligibility, and pronunciation teaching techniques, with a special focus on language rhythm
ES	**Elina Banzina.** Ph.D., es Profesora en el Departamento de Lenguas de la Escuela de Economía de Estocolmo, en Riga. Sus líneas de investigación incluyen la fonética inglesa, la percepción y la producción de la lengua hablada con acento extranjero, inteligibilidad y técnicas para la enseñanza de la pronunciación, con especial hincapié en el ritmo de la lengua.
IT	**Elina Banzina,** Ph.D., è Ricercatrice confermata presso Dipartimento di Lingue della Scuola di Economia di Stoccolma nella città di Riga. I suoi interessi di ricerca includono la fonetica inglese, la percezione e la produzione di parlato con accento straniero, intellegibilità e tecniche per l'insegnamento della pronuncia, con particolare enfasi sul ritmo della lingua.

Lynne E. Hewitt, Department of Communication Sciences and Disorders, Bowling Green State University

lhewitt@bgsu.edu

EN	**Lynne Hewitt,** Ph.D., CCC-SLP, is an Associate Professor and Chair of the Department of Communication Sciences and Disorders, Bowling Green State University, Bowling Green, OH. Her undergraduate training was in English and linguistics. She is a certified and licensed speech-language pathologist with clinical and research interests in language assessment and intervention.
ES	**Lynne Hewitt,** PhD., CCC-SLP (Certificado de competencia clínica en patologías del lenguaje), es Profesora y Directora del Departamento de Ciencias y Trastornos de la Comunicación en la Bowling Green State University en Bowling Green, en Ohio. Es licenciada en Inglés y Lingüística y posee certificación y licencia de patóloga del lenguaje. Sus intereses de investigación abarcan también la evaluación y la intervención lingüística.
IT	**Lynne Hewitt,** Ph.D., CCC-SLP (Certificato di competenza clinica in patologie del linguaggio), è Professore Associato e capo del Dipartimento di Scienze e Disordini della Comunicazione, alla Bowling Green State University sita a Bowling Green, in Ohio. Si è laurata in Inglese e Linguistica ed ha la certificazione e la licenza come patologa del linguaggio. I suoi interessi di ricerca riguardano anche la valutazione e l'intervento linguistici.

Laura C. Dilley, Department of Communicative Sciences and Disorders, Michigan State University

ldilley@msu.edu

EN	**Laura Dilley,** Ph.D, is an Assistant Professor in the Department of Communicative Sciences and Disorders, Michigan State University. Her research is in the area of phonetics, psycholinguistics, and phonology, particularly with respect to speech prosody.
ES	**Laura Dilley** Ph.D., es Profesora en el Departamento de Ciencias y Trastornos de la Comunicación en la Michigan State University. Su investigación se centra en el area de la fonética, psicolingüística y fonología, con especial interés para la prosodia de la lengua hablada.
IT	**Laura Dilley,** Ph.D, è Ricercatrice confermata presso il Dipartimento di Scienze e Disordini della Comunicazione, alla Michigan State University. I suoi interessi di ricerca sono nell'area della fonetica, della psicolinguistica e della fonologia, con particolare enfasi sulla prosodia del parlato.