Computer-Enhanced and Mobile-Assisted Language Learning: Emerging Issues and Trends

Combining the body and mobile technology to teach English pronunciation

Combining the body and mobile technology to teach English pronunciation

Felicia Zhang

University of Canberra, Australia

ABSTRACT

In Australia, postgraduate international student enrolments have risen 81 per cent since 2002. In 2008, 74 per cent of international enrolments were in Business/Management, Computer Science and Engineering-related courses (Davies, 2010). Most international students are from Asia, and the vast majority of those from mainland China. A survey of employers in 2006 found that employers rated interpersonal and communications skills (written, oral, listening) far above “qualifications” and “previous employment” as the most important selection criteria when hiring graduates (Davies, 2010). In fact, research on the IELTS levels of exiting Master’s degree students suggests that international students at Australian universities tended to graduate with worse spoken English skills than when they first entered.

The above research has identified a deficiency in the communication skills of international students, crucial in seeking employment. This chapter proposes various ways of helping committed international students on campus to improve their spoken English language by combining the use of the students’ bodies with mobile technology.

INTRODUCTION

According to Davies (2010), in Australia, postgraduate international student enrolments have risen 81 per cent since 2002. In 2008, 74 per cent of international enrolments were in Business/Management, Computer Science and Engineering-related courses. Most international students are from Asia, and the vast majority of those from mainland China. International students from China have been increasing 8 per cent since 2007. A worrying development is the decline in the employment success of international students, attributed largely to the inadequate English language skills of graduating students. In fact, research on the IELTS levels of exiting Masters students suggests that international students at Australian universities tended to graduate with worse spoken English skills than when they first entered (Birrell, 2006). As reported in Davis (2010), a survey of employers in 2006 found that employers rated interpersonal and communication skills (written, oral, listening) far above “qualifications” and “previous employment” as the most important selection criteria when hiring graduates (57.5 per cent, 35.4 per cent and 27.6 per cent respectively).

The above research has identified a deficiency in the communication skills of international students. One of the communication skills is obviously spoken English. However, in English language programs, spoken communication skills such as pronunciation has been marginalized for many years even in communicative language teaching (CLT) programs (Gatbonton, E. & Segalowitz, 2005). This is further compounded by many English language teachers’ lack of confidence in teaching the skill (Breitkreutz, Derwing, & Rossiter, 2001; Macdonald, 2002) despite the abundance of teaching materials available. Furthermore, the debate between what aspects of English pronunciation are more important to teach and what level of competency L2 learners should realistically aim for are still going on. For instance, research by Derwing and Munro (1997) and Derwing, Munro and Wiebe (1997; T. M. Derwing, Munro, & Wiebe, 1998) suggest that although second-language users are often able to modify their pronunciation to the extent that native listeners find their productions significantly easier to understand, total elimination of an accent is not a realistic goal. Otlowski (1998) and Fraser (1999) concur with much of the current research and suggested that the goals of pronunciation

teaching should be "developing functional intelligibility, communicability, increased self-confidence, the development of speech monitoring abilities and speech modification strategies for use beyond the classroom" (p. 3).

While the goal of pronunciation teaching and learning has moved from that of native-like pronunciation (as in the audio-lingual approach), to intelligibility (Celce-Murcia, Brinton, & Goodwin, 1996; Kenworthy, 1987; Morley, 1994), the comments of teachers on the coal face in Macdonald's study (2002) indicate that what intelligibility means and how it is measured are not clear. Intelligibility is therefore defined as the extent to which speech signal produced by the speaker can be identified by the listener as the words the speaker intended to produce. This definition is similar to that used by Field (2005, p. 401): "the extent to which the acoustic-phonetic content of the message is recognizable by a listener", and to the term 'phonological intelligibility' used by Jenkins (2002, p. 86). Intelligibility is a complex construct that is also tangled up with different views, personalities and experiences. Nevertheless, it is a concept pivotal to examination of learner goals and assessment for pronunciation.

With regards to what aspect of English pronunciation to focus on, i.e. segmental or suprasegmentals, some studies of international teaching assistants' (ITAs) speech found correlations between using suprasegmentals accurately and speaker intelligibility. Tyler, Jeffries, and Davies (1988) studied the discourse of ITAs whom undergraduates perceived as disorganized and unfocused. They found that these ITAs used too many pauses, too many primary stresses per message unit, and inappropriate falling intonation. The study conducted by Gallego (1990) also found that communication breakdowns were usually caused by pronunciation errors in word stress in ITAs' speech. Other research has suggested that a number of suprasegmental features such as wave duration and wave peak amplitude (Constantinou, 1993), word stress (Bansal, 1969), and the interstress interval (Anderson, 1993) affect intelligibility. Jenkins (2002) examined communication breakdowns of nonnative to nonnative speech, and found "a combination of phonological errors which caused the most serious problems in [her] data: misplaced tonic (nuclear) stress along with a consonant substitution within the wrongly stressed word" (p. 89). Morley (1991) and Fraser (1999) further call for a learner perspective in English pronunciation teaching through supporting a learner-centered approach that involves authentic tasks and the use of peers and groups for interaction and feedback to help learners be critical listeners and develop the ability to notice and repair their own and others' errors. Schmidt (1995), in a comprehensive survey of psychological studies, has also argued that second-language (L2) learners need to have their attention drawn to specific characteristics of a language if they are to make changes in their own productions. This chapter proposes such an approach in which a number of computer enhanced technologies including mobile technology are also called into service.

LITERATURE REVIEW

Correcting non-native speakers' English is not easy to accomplish. First of all, it is necessary to identify what errors L2 learners of English from different nationalities are making. Then, strategies for correction are needed, and because language use is essentially individualistic, somehow, the L2 learners of English need to know how to correct the errors themselves (H. Fraser, 1999).

In terms of identifying errors made by L2 learners of English from different nationalities, extensive research in linguistics, contrastive analysis, applied linguistics and spoken language technology for education (SLATE) have provided many useful clues (Ehsani & Knodt, 1998). Books describing L2 learner errors in English detail L2 learners' errors from different nationalities (Swan & Smith, 2001). However, pronunciation teaching is still rarely integrated into mainstream English language classes and, at best, occupies a marginalized position within the curriculum (Macdonald, 2002).

Recognizing the marginalization of pronunciation classes alerts us to the gap between research literature on pronunciation and classroom realities. For instance, many of the practices which were supposedly evidence-based, arose from research done in linguistics and SLATE. The research was usually carried out under controlled experimental conditions, for instance, selecting L2 learners from the same nationalities. Interesting as this may be, English classroom reality in Australia is that classes are likely to be composed of multiple nationalities, and this would cast doubt on the applicability of research findings obtained under controlled experimental conditions.

Macdonald's research on reasons why ESL teachers in Australia were reluctant to teach pronunciation also provided us with some reasons why English pronunciation teaching has been marginalized. He also made a number of recommendations (Macdonald, 2002). The two that will be discussed in details in this chapter are (i) how to provide teachers with guidance to teach students to set goals for themselves; and (ii) how to redefine the teacher's role with regard to pronunciation as that of a speech coach responsible for monitoring student speech and encouraging student self-monitoring.

The Role of Technology in Teaching English Pronunciation

Computer technology plays a central role in teaching English pronunciation. One is in its ability to assess non-native pronunciation and the other is through the use of computer assisted language learning programs to teach English pronunciation. Central to both is the development of automatic speech recognition (ASR) technology.

Computer Pronunciation Teaching Program

ASR has been used to construct computer assisted pronunciation teaching programs (CAPT) for a long time. Early ASR-based software programs, such as Talk to Me (Auralog, 1995), the Tell Me More Series (Auralog, 2000), Triple- Play Plus (Mackey & Choi, 1998), New Dynamic English (DynEd, 1997), English Discoveries (Edusoft, 1998), and See it, Hear It, SAY IT! (CPI, 1997) adopted template-based recognition systems which perform pattern matching using dynamic programming or other time normalization techniques (Dalby, J. & Kewley-Port, 1999). Neri, Cucchiarini, Strik and Bove (2002) criticized the graphical wave forms presented in products such as Talk to Me and Tell Me More because they look flashy to buyers but do not give meaningful feedback to users.

Even though, ASR technology has improved a great deal in the last ten years, most ASR engineers admit that the current accuracy level for a large vocabulary unit of speech (e.g., the sentence) remains less than 90%. For instance, Dragon's Naturally Speaking or IBM's ViaVoice, for example, show a baseline recognition accuracy of only 60% to 80%, depending upon accent, background noise, type of utterance, etc. (Kim, 2006). As cited in Kim (2006), "more expensive systems, Subarashii (Bernstein, J., Najmi, A., & F., 1999), EduSpeak (Franco, Neumeyer, Kim, & Ronen, 2001), Phonepass (R. Hinks, 2001), ISLE Project (Menzel, Herron, Morton, Bomaventura, & Howarth, 2001) and RAD (CSLU, 2003) are reported to outperform these two".

These ASR driven programs use the Hidden Markov Model (HMM) algorithms to deal with large units of speech (Ehsani & Knodt, 1998). That is, a speech recognizer based on HMM computes how close the phonemes of a spoken input are to a corresponding model, based on probability theory. High likelihood represents good pronunciation; low likelihood represents poor pronunciation (R. Hinks, 2003; Larocca, Moagan, & Bellinger, 1991).

More recent ASR programs that have adopted HMM such as FluSpeak (IBM ViaVoice recognizer) also measures intonation. However, as Kim (2006) pointed out since intonation refers to low or high pitches in a sentence, it is extremely difficult to register intonation using only a sound wave. This is because the value of pitch in a high tone is small (narrow), whereas that in a low tone is large (wide). An additional problem arises in that the value of pitches varies depending on the individual speaker. Thus, it is difficult to determine the degree of correct intonation for different people based on the absolute value of pitch in a sample.

Pronunciation Assessment

Since the 1990s Automatic Speech Recognition technology has been used firstly to assess non-native pronunciation. Zechner, Higgins, Lawless, Futagi, Ohls, & Ivanov (2009) have devised an automatic scoring method for use in rating the TOEFL iBT Practice Online product. They use a recognizer trained on a combination of American native (Broadcast News) and non-native speech that produces word identity, timing and confidence scores. Scoring of a non-native speaker's English speech is based on best fit to their human gold standard, i.e. the judgment of a set of human listeners who are often experts in some aspect of language learning, such as classroom teachers or heads of national testing groups. Other similar speech assessment tools also exist commercially such as the Pearson Test of English (Academic) (Pearson, 2009). Pronunciation

assessment of this kind is based on the judgment of speech on its natural flow (use of pauses, rhythm, use of pitch, etc.), overall correctness of articulation and other criteria that give the listener the impression of fluent communication. In tests of this kind, the mastery of suprasegmental characteristics of English is essential in L2 speakers of English being perceived as intelligible speakers of English.

The CAPT systems reviewed above are less likely to provide us with reliable measures the suprasegmental aspects of English pronunciation and yet L2 speakers are increasing being tested on the suprasegmental aspects of English pronunciation. This makes them unsuitable as the sole means of teaching English pronunciation. However, many institutions and teachers will use commercial CAPT programs to 'take care' of the teaching of English pronunciation. No matter how advanced or accurate CAPT systems are, it is difficult to build in the feedback, instructions that students can act upon.

In this chapter, I would like to propose a method of teaching English pronunciation which teaches L2 learners how to set learner goals (Macdonald, 2002, recommendation 2) and learn to self-monitor and how to train English teachers to integrate the teaching of English pronunciation in their classroom and develop the ability to monitor L2 students' pronunciation (Macdonald, 2002, recommendation 3) with and without the use of CAPT technology. The technology used in this approach includes cd roms, speech analysis tool and the Smart Pulse Pen by Livescribe (<http://www.livescribe.com>).

TOWARDS AN EMBODIED APPROACH TO THE TEACHING OF ENGLISH PRONUNCIATION

In Australia, a large majority of L2 students in ELICOS centers want to undertake degree studies in universities. Therefore, they need to fulfill entry requirements such as that for English proficiency. The most commonly used test of English proficiency is the IELTS test. In the IELTS test, the teaching of speaking skill in English comprises of the teaching of fluency and coherence, lexical resources, grammar range and accuracy and pronunciation according to the IELTS Band Descriptors (University of Cambridge ESOL Examination Board, 2010). Fluency is a commonly used notion in foreign language teaching and yet it is a concept that is difficult to define precisely. Even without its precise definition, fluency continuous to be one of the aspects of spoken English tested on most common descriptors of oral performance in English. the Concise Oxford Dictionary defined fluency in a language as "smooth, rapid, effortless use of language".

This view of fluency has been accepted in the definition of fluency according to, e.g. IELTS speaking skill band descriptors which states that at Band 7:

- speaks at length without noticeable effort or loss of coherence
- may demonstrate language-related hesitation at times, or some repetition and/or self-correction
- uses a range of connectives and discourse markers with some flexibility'

Intelligibility only appears in the descriptor at Band 8 and 9 under pronunciation in the form of 'Pronunciation is effortless to understand.' and 'Pronunciation is easy to understand throughout, L1 accent has minimal effect on intelligibility.'

Since in most institutions, the minimum entry requirement for university students is IELTS overall score of 6.5 which includes the speaking skill, then we need to integrate both fluency and intelligibility activities in pronunciation teaching. What kind of language do we select to teach?

Going back to the IELTS band description, exactly how can we develop in L2 learners' the ability to speak at length without noticeable effort or loss of coherence? To answer this question, we need to first discuss what kind of language we need to select to teach and then investigate how to teach these target features.

Many studies have been conducted to investigate, from a listener's perspective, what aspects of L2 learners' English affect comprehension in native speakers of English. Anderson-Hsieh and Koehler (1988) found that prosodic deviance may affect comprehension more adversely than segmental deviance. Zielinski (2006) provides some discussion of other studies that have attempted to identify features contributing to intelligibility. Romova, Smith and Neville-Barton's work (2008) also correlates with that of Anderson-Hsieh and Koehler (1988). However, the issue of what factors impact most on listeners' intelligibility is far from conclusive (T. M.

Derwing & Munro, 1997{Munro, 1995a #398}). Recent work on a group of international teaching assistants (ITAs) (Hahn, 2004) found that primary stress or sentence stress is a major factor influencing the intelligibility of ITAs' speech and that native English speakers could recall significantly more content and could process the discourse more easily if these characteristics of the language are correct. This also led to a positive evaluation of the L2 speaker. In spoken conversation, intonation and stress information not only helps listeners to locate phrase boundaries and word emphasis, but also to identify the pragmatic thrust of the utterance (e.g., interrogative vs. declarative).

Grosjean and Gee (1987) argued that the rhythmic properties of the speech signal are important for listeners and hypothesized that listeners use strong syllables to initiate a search for words in that minds. Research on the process by which native English listeners derive a sequence of words from a stream of connected speech indicated that the rhythmic properties of the speech signal are also important to native speakers of English as they draw on these properties to divide continuous speech into individual words and to recognize what the individual words are (Cutler & Butterfield, 1992).

The research reviewed above is in general agreement that sentence and word stress are important for intelligibility and that errors made in word stress within sentences were responsible for causing misunderstanding in native speakers. However this should not be interpreted as that in pronunciation teaching, we should start with words. Since sentences are carriers of the rhythm of a language, the lowest teaching unit should be sentences rather than words. Then the question is 'which sentences?'

Many researchers found that formulaic sequences constitute a large proportion of spoken discourse (Schmitt & Carter, 2004). Altenberg (1998) found that 80% of the words in the London–Lund corpus of spoken English form part of formulaic sequences, remarking that "what is perhaps the most striking impression that emerges (. . .) is the pervasive and varied character of conventionalized language in spoken discourse (. . .) from entire utterances operating at discourse level to smaller units acting as single words and phrases" (p. 121). So which formulaic sequences?

Gatbonton and Segalowitz (2005) argue that since the formulaic nature of an utterance cannot be fully divorced from the contexts in which it occurs, many more such utterances can be found if searched for in specific pragmatic contexts (Coulmas, 1981) and communicative situations (Kecskes, 2002; Read & Nation, 2004).

Why should formulaic utterances be the focus of English pronunciation teaching and can the teaching of these utterances lead to grammatical competence? This question concerns the debate on whether learning a L2 consists of acquiring rules (rule-based learning) or simply mastering specific exemplars of language (Hopper, 1998; Kemmer & Barlow, 1999). Some have suggested that language acquirers do not learn predetermined sets of rules (Hopper, 1998, p. 156) but learn particular instances of usage and construct commonalities from these instances as they go along (Goldberg, 1998; Healy & Schneider, 1998) (Goldberg, 1998, p. 209; Healy et al., 1998, p. 18; see also Kemmer & Barlow, 1999). Skehan (1996) suggests that both rules and specific utterances are learned either simultaneously or successively (exemplar-based learning occurring in the early stages followed by rule-based learning). Ellis (2002) identifies frequency-based exemplar learning as a major force in second language learning.

Levelt (1989), writing on speech production in native speakers, emphasizes that fluency in spoken language production in first language requires automaticity. What gets automatized are frequently used exemplars. It seems likely that fluency is greatly enhanced by the control of large numbers of frequently used exemplars. These prefabricated pieces are often strung together in a way appropriate to the communicative situation, allowing the speaker's energy or attention linked with single lexical units in the speech run to be freed up to plan larger stretches of speech. In many familiar communicative situations, concepts and speech acts can be expressed formulaically. So when communicating in such situations, if a speaker can pull these formulas readily from memory, that is, if these utterances are automatized, fluency is enhanced. Nattinger and De Carrico (1992) makes a strong case for teaching prefabricated units or patterns because they reduce the processing effort. Wood (2007) suggested that this automatic production of formulaic utterances reduces the amount of planning, processing and encoding needed within clauses thus giving 'the speaker time to pay attention to the multitude of other tasks necessary while speaking, such as generating specific lexical items, planning the next unit of discourse, syntactic processing of novel pieces and so on' (Wood, 2007).

If fluency is an observable indication of a speaker's control of formulaic utterances, what features of non-native speakers' speech are indications of fluency speak?

A significant amount of research (Cucchiari, Strik, & Boves, 2000; Freed, 1995; Lennon, 1990; Riggensbach, 1991) found that native judges may usually apply a blending of linguistic proficiency and effortless production in their judgment. In Freed's study, some native speakers' judgment were influenced by accent and intonation but for most judges, both rate of speech and number of pauses were major factors influencing their judgment of fluency (Freed, 1995). However, there seems to be a growing consensus in more recent research that fluency cannot be reduced to speed of delivery and that articulation rate has less impact on our perception of fluency than the length, the nature and the location of pauses in the utterance. Natural pauses, allowing breathing space, usually occur at some clause junctions or after groups of words forming a semantic unit. Pauses appearing at places other than these are judged as hesitations, revealing either lexical or morphological uncertainty. These hesitations may be either simply a silent gap or marked by non-lexical fillers such as "uh", "um", sound stretches (or drawls on words) or lexical fillers with no semantic information (such as "you know", "I mean"). Cucchiari et al's study confirmed that becoming fluent therefore is not just about speaking faster (articulation rate), but about pausing less often and pausing at the appropriate junctures in an utterance. What appears significant from research in this area is:

- the frequency of pauses rather than the length,
- the length of run (the number of syllables between pauses),
- the place of pauses in an utterance,
- the transfer (or not) of pausing pattern from L1 to L2.

Given that we need to focus on sentence stress, rhythm of English, increasing the length of run spoken by L2 students and the development of an ability to place pauses in the right places, a teaching method that involves the whole body in the process of learning is necessary.

THE PILOT STUDY

In the pilot study, 4 non-native speakers (NNS) of English's read speech were first analyzed against speech read by 6 young cultured female native speakers of English taken from the Australian National Database of Spoken Language (ANDOSL) database. Out of the four NNS, one is from Japan and three others are from Taiwan. After the initial analysis, the Japanese speaker underwent an intervention using the Somatic-Enhanced Approach (SEA) to correct her pronunciation. The intervention consisted of 10 hours of face-to-face contact. The Japanese student reported improved spoken English and improved listening ability in English.

The ANDOSL database contains phonetically-rich read material and spontaneously spoken culturally relevant material. 10 sentences read by both groups of speakers were recorded and analyzed according to rate of speech (ros), number of pauses, total duration of pauses and mean length of pauses (mlp). Rate of speech (ros) was defined as the number of syllables divided by the total duration of speech including sentence-internal pauses. Number of pauses (#p) is defined as number of sentence-internal pauses of no less than 0.2s. Total duration of pauses (tdp) is defined as the total duration of all sentence-internal pauses of no less than 0.2s. Mean length of pauses (mlp) is defined as the mean length of all sentence-internal pauses of no less than 0.2s. The ten sentences used were:

- S1: **The** price range **is** **smaller** **than** any **of** **us** expected
- S2: They asked if I **wanted** **to** come **along** on **the** barge trip
- S3: Amongst **her** friends she **was** **considered** beautiful
- S4: **The** smell **of** **the** freshly ground coffee **never** fails to entice me **into** **the** shop
- S5: I'm **often** **perplexed** by rapid **advances** in state **of** the art technology
- S6: John **could** lend him **the** latest draft **of** his work
- S7: **From** forty love **the** score **was** now deuce **and** **the** crowd grew tense
- S8: **The** Presbyterian minister managed to curb **the** drinking habits **of** **the** loitering youths
- S9: **The** bulb blew when he switched on **the** light
- S10: It **is** futile to offer any further **resistance**

Words that should be pronounced as schwa have been bolded.

In order to make sure that native speakers would pronounce the same words as 'schwa', these ten sentences were subjected to 'schwa' analysis by two native speakers. One native speaker underlined 52 words in the above ten sentences as schwa. The second native speaker underlined 34 'schwa' elements. These are used as a benchmark to measure whether our Japanese subject can pronounce words with 'schwa' or not.

Table 1. Characteristics of speech of the six native speakers from the ANDOSL database

Native Speakers	Average ROS	#p	tdp	mlp
Andosl speaker 1	4.17	0	0	0
Andosl speaker 2	4.30	1	0.37	0.37
Andosl speaker 3	4.70	1	0.22	0.22
Andosl speaker 4	6.20	1	0.39	0.39
Andosl speaker 5	5.94	4	0.90	0.23
Andosl speaker 6	5.67	0	0	0
average	5.16	1.17	0.31	0.27

Key:

Ros= the number of syllables divided by the total duration of speech including sentence-internal pauses.

#p= number of sentence-internal pauses of no less than 0.2s.

Tdp= the total duration of all sentence-internal pauses of no less than 0.2s.

Mlp= the mean length of all sentence-internal pauses of no less than 0.2s.

Table 2: Characteristics of speech of the three non-native speakers

L2 Speakers of English	Average ROS	#p	tdp	mlp	IELTS score
Japanese (Aiko)	2.87	7	3.04	0.43	6.5
Chinese 1	3.48	2	1.32	0.66	7
Chinese 2	3.21	6	3.22	0.54	7.5
Chinese 3	3.86	3	1.51	0.50	8
average	3.36	4.5	2.27	0.53	7.25

Key:

Ros= the number of syllables divided by the total duration of speech including sentence-internal pauses.

#p= number of sentence-internal pauses of no less than 0.2s.

tdp= the total duration of all sentence-internal pauses of no less than 0.2s.

mlp= the mean length of all sentence-internal pauses of no less than 0.2s.

According to table 1, the six female native speakers (ns) of Australian English spoke on average 5.16 syllables per second; most of them (except one person) made at least one pause in the ten sentences. Speaker made 4 pauses of shorter than average length. The average length of a pause for the group was 0.27s.

Comparatively speaking, according to table 2, the Japanese non-native speaker (nns) spoke only about 2.87 syllables per second and made as many as 7 pauses with a total duration of pauses of 3.04. The mean length of her pauses was 0.43 s. For the other NNS students, they spoken faster on average with 3.36 syllables per second and made a total of 18 pauses (on average 4.5 pauses per person) with a mean length of pauses as 0.53s.

For all NNS involved in the pilot project, the positions of the pauses are all internal to words rather than at the appropriate junction. The causes of these pausing errors were the unfamiliarity with words. The words these L2 learners had problems with were: S1: expected; S2: asked S3: amongst, S4: entice; S5: perplexed and advances; S7: deuce; S8: Presbyterian, loitering and in S10: futile

These findings concur with previous research on temporal variables of fluent speech which demonstrated that dysfluency in NNS' English speech is marked by slower rate of speech, frequency of pauses and the unusual location of the pauses.

Focusing on the Japanese speaker (Aiko, pseudonym), reading the same ten sentences, she spoken a lot slower and made almost as many pauses as the Chinese L2 speakers. Analysis of Aiko's reading of the ten sentences suggests that she did not know how to link words together (liaison), had problems with 'schwa' and had problems with the end of words such as 'asked' and 'wanted', 'along/arong) in sentence 2. She also seemed to be influenced by the orthography of the sentences because she was reading the sentences.

For instance, in sentence 1, she had difficulty with the string 'any of us' and in sentence 9, she did not link up 'switched' with 'on'. In order to highlight the importance of linking in English, I chose the following poem to make her first perceive the importance of linking in English and then learn how to physically produce the linking in English. Though some researchers would argue that for pronunciation teaching to be effective teachers should use materials in class that are as close as possible to the speech learners would use outside the teaching situation (2001), in SEA the use of poems (as illustrated here) was only used to make students' aware of certain characteristics of English such as liaison, the stress-time nature of the language and so on. In using SEA to correct L2 learners' English, language that is similar to what the students would encounter outside the teaching situation would also be used.

The poem chosen was the popular nursery rhyme:

Cinderella dressed in yellow,
Went up town to meet her fellow,
She walked ed so slow, she met her beau,
He took her to a picture show,
How many kisses did he give her,
One, two, three, four.

(Taken from Celce-Murcia, Brinton, D. M. and Goodwin, J. M (1996, p. 300))

First of all, I recorded the poem for her to transcribe on a computer and asked her to do this as listening homework. When I saw her the following week, I found out that she could not transcribe the linked phrase 'went up town'. I did not immediately show her what the words were in the original poem. What I did instead was to follow the following steps to teach her the poem:

Step 1: is the relaxation step of SEA (Step 1) is not unlike the guided imagery activities advocated in psychology such as neurolinguistic programming (NLP)-a set of pedagogical beliefs rooted in Counseling-Learning theory and Gestalt therapy(Richard Bandler & John Grinder, 1975; Richard Bandler & John Grinder, 1975; Grinder & Bandler, 1979).

Step 2:

I hummed the first phrase 'Cinderella dressed in yellow' in 4/4 beat and asked Aiko to walk in a circle with me and repeat the hummed sentence in the same 4/4 beat with me. This was done 5 times. This was used to highlight the syllable structure of the sentence.

Step 3: I then clapped and hummed to the same phrase by clapping to the rhythm in 4/4 beat.

Step 4: I then continue to clap to the rhythm of the sentence in 4/4 beat but when I came to the linked words 'dressed in' I stamped my foot to emphasize the link between 'ed and in'. This again was done 5 times.

Step 5: In this step, I would say the phrase while instructing Aiko to only mouth the phrase but not to sound it out. Again this was done 5 times.

Step 6: In this step, I asked Aiko to still continue to walk in circles and then repeat the target sentence with me. This is the first time she was hearing the full sentence in its original rhythmic context.

By the time Aiko finishes step 3, she would already have uttered the sentence with the correct rhythm, intonation and stress, the use of gesture in step 4 then 'anchors' the previous learning for future recall so that such learning can be readily activated by students themselves at the point of need. Throughout the procedure, Aiko also had plenty of opportunities to reflect on the sentences she was learning. In teaching the third line 'She walked so slow, she met her beau', recalling that she made an error with the word 'asked' in sentence 2, I

used the gesture of stretching a string while walking to emphasise the length of the word 'walked' (walkt) in this particular line.

It was found that after two hours of such training, the L2 student's awareness of English rhythm and stress patterns was heightened. She also reported that her listening skill in English also improved. I chose to use the procedure to ask Aiko to directly experience the different characteristics of the sentence through her perceptual system rather than through the medium of the written word. This was deliberately chosen to remove the influence of English orthography.

ROLES OF THE SPEECH ANALYSIS TOOL (SPTOOL) AND SMART PULSE PENS

The use of Praat (open-source acoustic analysis software) to provide feedback in pronunciation classes promotes autonomous learning in a field that has had to rely on native listener judgments (Wilson, 2008). Wilson (2008) further demonstrated the use of Praat to offer feedback on Japanese L2 students of English's duration, pitch, and intensity of sounds although the paper did not contain data that illustrated the transfer effect of these activities on learning. Praat is the same software that is used to analyze data for this chapter. However, further validation of using Praat for language learning is needed.

Sptool and Data CD-Rom

The author, however, has been using Sptool (F Zhang, 2003) which was developed and tested with numerous groups of foreign language learners (Buranapatana & Zhang, 2007; F Zhang, 2005; F Zhang & Wagner, 2005). It was designed to give learners the opportunity to listen and produce their own speech and then compare it to the native speaker model. Such comparison increases the probability of changing the ways they both hear and produce the target language. By contrast to other ASR-based CAPT system, Sptool (F Zhang, 2003) is designed to offer feedback that is non-judgmental, i.e. without a scoring system especially when the limitation of the current ASR systems prevent feedback on suprasegmentals to be measured accurately (Kim, 2006). Combined with the procedures in SEA, it allows students to explore and reflect during the process of learning, and not just at the end of the learning process. Recent research (Buranapatana, in press) suggests that L2 students of Thai used Sptool as a teacher providing modeling and feedback to students thus making them more confident in learning and speaking Thai.

In the pilot study described in this chapter, another use was found for Sptool. During the intervention phase of the pilot study, another native speaking ESL teacher was also recruited to teach. Originally the intention was to train him to teach Aiko using SEA but this was found to be difficult as the native speaking teacher was not familiar with the rhythm of his own language and kept on changing the pitch of the target sentence. For example, in teaching the following role play:

Sales person: Buy now, pay later
Interested buyer: Wait a minute, let me think.

While I introduced the sentence 'Buy now, pay later' according to the following pitch curve with the expressive aim of teaching stress on 'now,' seen in Figure 1 and Figure 2.

Figure 1. Version one of the phrase "Buy now, pay later."

The native speaker kept on changing the phrase to the following curve where 'now' actually goes up rather than down.

Figure 2. Version two of the same phrase “Buy now, pay later.”

Because in English, the same sentence can be expressed in different ways depending on what the speaker's intention is, version 2 is of course also acceptable. But if the expressed aim of this sentence is to teach sentence stress with the accompanying gesture of stamping the foot on the ground, then it is essential that each time (as we walked around in circles) the sentence is said the same way. Otherwise students will become confused.

In order to keep the sentence the same, I recorded the sentence on Sptool and used that to remind the native speaking teacher how the sentence should be said. This proved to be very useful guidance for the teacher. Since every sentence can be extracted with Sptool, students can also revise the same sentences covered in class in the comfort of their home.

The pilot project on using SEA to teach English pronunciation demonstrated the fact that native ESL/EFL speakers do not usually have a 'feel' for the rhythm of their own mother tongue. In fact, a period of training is necessary for them to feel comfortable with the approach. Before they teach, they need to figure out what the rhythm, stress, intonation patterns of the sentences they want to teach are. Sptool can help them in this preparation and help them to keep to the same intonation pattern each time they teach the sentence during the session.

Fraser (2001) also pointed out that the aim of pronunciation lessons is 'to give the learners information and guidance that they can act upon to change the way they speak' and therefore it is important to set up the metalinguistic communication between the teacher and the students. The use of visual cues through Sptool is part of such metalinguistic communication. In her research, Fraser (2001) used annotated spelling as visual cues. In SEA, visual cues are provided through gestures as well as through Sptool. The understanding and use of both types of visual cues have been built up in class between the teacher and the students. This way, L2 students would learn how to interpret the visual cues with the guidance of the teacher. Through the use of

Sptool, L2 students also learn critical listening skills (Fraser, 2001) which enable them to monitor their own progress in English pronunciation.

Livescribe Smart Pulse Pen

Another useful tool for language learning is the Smart Pulse pen (<http://www.livescribe.com/en-au/smartpen/pulse/>). Two of the main components of the Livescribe platform are:

- Smartpen: a Montblanc-size computer with advanced processing power, audio/visual feedback, and substantial memory for handwriting capture, audio recording, and applications;
- Dot Paper with Dot Positioning System (DPS): technology that enables interactive, “live” documents using plain paper printed with micro-dots. (<http://www.livescribe.com/en-au/smartpen/pulse/>).

In my lectures, I use Livescribe’s first key application is “Paper Replay” to take notes during a discussion or lecture. The Smartpen also records the conversation and digitizes the handwriting, automatically synching the ink and audio. By later tapping the ink, the Smartpen replays the conversation from the exact moment the note was written. I also upload the notes and audio to my PC and later uploaded to the Livescribe website for my students to access. All notes and audio can be replayed, saved, searched and sent through Livescribe desktop.

This year in my Chinese class, I also used the Smartpen in combination the dot paper to assist a student with disability to keep up with the Chinese unit he was doing. I printed the entire teaching material onto the dot paper that works with the Smartpen and recorded each sentence on the page in Chinese. This way when the student taps on the sentence, he/she can hear the audio file for the sentence. This student was helped greatly by the Smartpen and the textbook even though he was absent from tutorials due to circumstances beyond his control. An interview with the student suggested:

I guess I just learn all the characters, quite difficult. Using the pen and I can listen to the characters as well. I can relate to them. The pen has all the sections recorded on it so I turn the pen on and listen to it, I relate sounds to character. There is no need to be in front of the computer. I do an hour each day. I prefer the pen to the cd, because it is more accessible and portable.

Smartpens are currently still too expensive for extensive experimentation but anecdotal reports suggest that packaging textbooks on dot paper and the Smartpen makes learning mobile and has the potential to increase time on task.

CONCLUSION AND DISCUSSION

This chapter outlines the SEA approach, which aims directly at promoting the teaching of English pronunciation by re-emphasizing the central role of the body in initiating and sustaining the automatization of the language learned in a manner compatible with a communicative approach to language teaching. It was argued that for this approach to succeed in the goal of automatizing the language learned the target language items (sentences rather than words) need to be experienced and practiced through the L2 learners’ perceptual systems. Fraser pointed out (Helen. Fraser, 2001) that:

The subconscious concepts that actually drive our understanding and behavior can only be learned or altered through experience and practice. A teacher can only ever be a facilitator in this process – the learning must be done by the learners themselves. Explicit teaching can only affect conscious concepts. This is often described in terms of a difference between knowing that and knowing how. For the conscious concepts of knowing that to actually affect our behavior or knowing how, they need to filter down to the subconscious level.

The steps in SEA are illustrations of how ‘conscious concepts of knowing that’ (initially within the domain of the teacher’s knowledge) can filter down to the L2 students’ subconscious level through the use of movement and gesture consistently and repetitively throughout the classroom procedure. L2 students therefore, learn through guided experience which helps them build up appropriate subconscious concepts (Helen. Fraser, 2001). The build in repetition is also important because ‘from a psychological perspective, because of the high

consistency of situation-utterance correspondences across the repeated events, this repetition will lead to automaticity in both reception and production (Schneider & Chein, 2003).

The most important criteria for this approach to succeed is for it to be driven by evidence of errors collected from L2 learners' speech so that corrective measures applied are more targeted. In certain respects, many of the features contained in SEA are traditional, and teachers will easily recognize many of the illustrative activities as similar to ones they have developed or used. Many of the activities are similar to progressions in task-based activities as advocated by Willis (1996). Two things about SEA, however, are new. One is its focus on automaticity and the goal of promoting it primarily through perceptual and productive activities embodied in the classroom procedure. The other is the use of computer-enhanced technology and mobile technology such as Sptool and the Smart Pulse pen to enhance the provision of feedback to students.

It is argued that that SEA can be applied to the teaching of English pronunciation in order to help L2 learners of English to develop 'functional intelligibility, communicability, increased self-confidence, the development of speech monitoring abilities and speech modification strategies for use beyond the classroom' as advocated by Fraser (2001).

The methodological perspective presented here is not intended to supplant current CLT methodologies. In fact, it can be part of any CLT classrooms. Making the automatization of utterances a central goal means it is possible for a CLT methodology such as SEA to be designed to allow students to go home with things they can act upon (e.g., practice, compare and reflect on, and use essential language learned) with the help of the mobile learning tools. Experience with teaching Chinese as a foreign language in Australia suggests L2 students are no longer learning the way we used to do. They now learn through looking up dictionaries online, on their iPod touch, iPads or iPhones.

A book printed on Livescribe dot paper with the Smart pulse pen might look like a return to the book but its mobility and easy access to sound files mean that students can now learn through their ears without the interference of the written script in any language. Therefore, if the methodological principle is to train them to perceive certain sounds or differences between a particular sound in their mother tongue and in the target language (i.e. differentiating 'r' from 'l' in English for Japanese learners of English), then with the help of the Livescribe dot paper and the Smart pulse pen, we can conduct research into what kind of materials might be suitable for this kind of training. Such arrangement of teaching and learning materials has enormous potential for the future of foreign language learning.

Finally, SEA brings together insights and knowledge gathered from personal experience as well as from a rich body of research described in the literature review. The attempt is to make the approach testable by making the design criteria explicit and by founding it on basic principles concerning learning, memory, attention, and skill acquisition drawn from cognitive science. It is hoped future research with SEA will yield improved practical solutions for the teaching of English pronunciation.

REFERENCES

- Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 101-122). Oxford, UK: Clarendon.
- Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning*, 38(4), 561-613.
- Anderson, P. (1993). *The interstress interval as an indicator of perceived intelligibility among nonnative speakers of English*. Unpublished PhD thesis, Wichita State University.
- Bandler, R., & Grinder, J. (1975). *The structure of magic I: A book about language and therapy*. Palo Alto, CA: Science & Behavior Books.
- Bandler, R., & Grinder, J. (1975). *The structure of magic II: A book about communication and change*. Palo Alto, CA: Science & Behavior Books.
- Bansal, R. (1969). *The intelligibility of Indian English*. Hyderabad, India: Central Institute of English.
- Bernstein, J., Najmi, A., & Farzad, E. (1999). Subarashii: Encounters in Japanese spoken language education. *CALICO*, 16(3), 361-384.

- Birrell, B. (2006). Implications of low English standards among overseas students at Australian universities. *People and Place*, 14(4), 53.
- Breitkreutz, J. A., Derwing, T. M., & Rossiter, M. J. (2001). Pronunciation teaching practices in Canada. *Test Canada Journaurevue Test Du Canada*, 19(1), 51-61.
- Buranapatana, M., & Zhang, F. (2007). *Supporting learning flow through integrative technologies*. Paper presented at the 15th International Conference on Computers in Education (ICCE 2007), Hiroshima.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge, UK: Cambridge University Press.
- Constantinou, A. (1993). *Intelligibility of Mandarin speakers of English: Correlation of acoustic and perceptual measures*. Unpublished Masters thesis, California State University, Long Beach.
- Coulmas, F. (1981). Introduction: Conversational routine. In F. Coulmas (Ed.), *Conversational routine: Explorations in standardized communication situations and prepatterned speech* (pp. 1-18). The Hague, The Netherlands: Mouton.
- CSLU. (2003). *Toolkit*. Retrieved from <http://cslu.cse.ogi.edu/toolkit>.
- Cucchiaroni, C., Strik, H., & Boves, L. (2000). Quantitative assessment of second language learners' fluency by means of automatic speech recognition technology. *J. Acoust. Soc. Am*, 107(2), 989-999.
- Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, 31, 218-236.
- Dalby, J., & Kewley-Port, D. (1999). Explicit pronunciation training using automatic speech recognition. *CALICO*, 16(3), 425-445.
- Davies, M. (2010). International students and the permanent residency rort. *Quadrant Online*, 54(3).
- Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 19, 1-16.
- Derwing, T. M., Munro, M. J., & Wiebe, G. E. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3). doi: 10.1111/0023-8333.00047
- Ehsani, F., & Knodt, E. (1998). Speech technology in computer-aided language learning: Strengths and limitations of a new CALL paradigm. *Language Learning and Technology*, 2(1), 45-60.
- Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, 39(3), 399-423.
- Franco, H., Neumeyer, L., Kim, Y., & Ronen, O. (2001). Automatic pronunciation scoring for language instruction. *Proceedings of International Conference on Acoustics, Speech, and Signal Processing*, (pp. 1471-1474). Retrieved from http://www.speech.sri.com/people/hef/papers/icassp97_pronunciation.pdf
- Fraser, H. (1999). ESL pronunciation teaching: Could it be more effective? *Australian Language Matters*, 7(4), 7-8.
- Fraser, H. (2001). *Teaching pronunciation: A handbook for teachers and trainers*. Sydney, Australia: TAFE NSW Access Division.
- Freed, B. (1995). What makes us think that students who study abroad become fluent? In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-148). Amsterdam, The Netherlands: John Benjamins.
- Gallego, J. C. (1990). The intelligibility of three nonnative English speaking teaching assistants: An analysis of student-reported communication breakdowns. *Issues in Applied Linguistics*, 1(2), 219-237.
- Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 61(3), 325-353.
- Goldberg, A. (1998). Patterns of experience in patterns of language. In M. Tomasello (Ed.), *The new psychology of language* (pp. 203-219). Mahwah, NJ: Lawrence Erlbaum.
- Grinder, J., & Bandler, R. (1979). *Frogs into princes: Neuro linguistic programming*. Moab, UT: Real People Press.
- Grosjean, F., & Gee, J. P. (1987). Prosodic structure and spoken word recognition. *Cognition*, 25, 135-155.

- Hahn, L. (2004). Primary stress and intelligibility research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38(2), 201-223.
- Healy, A., Barshi, L., Crutcher, R. J., Tao, L., Rickard, T. C., Marmie, W. R., ... Bourne, L. E., Jr. (1998). Toward the improvement of training in foreign languages. In J. A. Healy & L. E. Bourne (Ed.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 3-35). Mahwah, NJ: Lawrence Erlbaum.
- Hinks, R. (2001). Using speech recognition to evaluate skills in spoken English. *Working Papers*, 49, 58-61. Lund University, Department of Linguistics.
- Hinks, R. (2003). Speech technologies for pronunciation feedback and evaluation. *ReCALL*, 15(1), 3-20.
- Hopper, P. J. (1998). Emergent grammar. In M. Tomasello (Ed.), *The new psychology of language* (pp. 155-175). Mahwah, NJ: Lawrence Erlbaum.
- Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics*, 23(1), 83-103.
- Kecskes, I. (2002). *Situation-bound utterances in L1 and L2*. Berlin, Germany: Mouton de Gruyter.
- Kemmer, S., & Barlow, M. (1999). A usage-based conception of language. In M. Barlow & S. Kemmer (Eds.), *Usage-based models of language* (pp. xii-xxiii). Stanford, CA: Center for the Study of Language and Information.
- Kenworthy, J. (1987). *Teaching English pronunciation*. London, UK: Longman.
- Kim, I.-S. (2006). Automatic speech recognition: Reliability and pedagogical implications for teaching pronunciation. *Educational Technology & Society*, 9(1), 322-334.
- Larocca, S. T., Moagan, J. J., & Bellinger, S. M. (1991). On the path to 2X learning: Exploring the possibilities of advanced speech recognition. *CALICO*, 16(3), 295-310.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40, 387-417.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Macdonald, S. (2002). Pronunciation - Views and practices of reluctant teachers. *Prospect* 17(3), 1-18.
- Menzel, M., Herron, D., Morton, R., Bomaventura, P., & Howarth, P. (2001). Interactive pronunciation training. *ReCALL*, 13(1), 67-78.
- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25(1), 51-74.
- Morley, J. (1994). A multidimensional curriculum design. In J. Morley (Ed.), *Pronunciation pedagogy and theory* (pp. 64-91). Alexandria, VA: TESOL.
- Nattinger, J. R., & De Carrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford, UK: Oxford University Press.
- Neri, A., C. Cucchiari, Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer Assisted Language Learning*, 15(5), 441-467.
- Otlowski, M. (1998). Pronunciation: What are the expectations. *I-TESL-J*, 4(1).
- Pearson, C. (2009). *Pearson test of English: Academic*. Retrieved from <http://pearsonpte.com/Pages/Home.aspx>
- Read, J., & Nation, P. (2004). Measurement of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 23-35). Amsterdam, The Netherlands: John Benjamins.
- Riggenbach, H. (1991). Towards an understanding of fluency: A microanalysis of nonnative speaker conversation. *Discourse Processes*, 14, 423-441.
- Romova, Z., Smith, J., & Neville-Barton, P. (2008). Can I change the way I speak? An exploration into pronunciation and fluency after three years of tertiary EAL study. *Prospect*, 23(3), 12-23.
- Schmidt, R. (1995). *Attention and awareness in foreign language learning*. (Technical Report No. 9). Hawaii: University of Hawaii Second Language Teaching and Curriculum Center.
- Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use* (pp. 1-22). Amsterdam, The Netherlands: John Benjamins.
- Schneider, W., & Chein, J. M. (2003). Controlled & automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science*, 27, 525-559.

- Swan, M., & Smith, B. (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge University Press.
- Tyler, A., Jeffries, A., & Davies, C. (1988). The effect of discourse structuring devices on listener perceptions of coherence in non-native university teachers' spoken discourse. *World Englishes*, 7(2), 101-110.
- University of Cambridge ESOL Examination Board. (2010). *IELTS speaking band descriptors* (public version). Retrieved from <https://www.teachers.cambridgeesol.org/ts/exams/academicenglish/ielts>
- Willis, J. (1996). *A framework for task-based learning*. London, UK: Longman.
- Wilson, I. (2008). *Using Praat and Moodle for teaching segmental and suprasegmental pronunciation*. Paper presented at the 3rd International WorldCALL Conference: Using Technologies for Language Learning (WorldCALL 2008).
- Wood, D. (2007). Measuring the link between formulaic sequences and speech fluency: Implications for the language classroom. *Contact*, 33(2), 97-117.
- Zechner, K., Higgins, D., Lawless, R., Futagi, Y., Ohls, S., & Ivanov, G. (2009). *Adapting the acoustic model of a speech recognizer for varied proficiency non-native spontaneous speech using read speech with language-specific pronunciation difficulty*. Paper presented at the Interspeech 2009, Brighton, U.K.
- Zhang, F. (2003). *Speech tool* [Software]. Canberra, Australia: University of Canberra, Australia.
- Zhang, F. (2005). Making feedback last: An integrated approach to feedback in language learning. In J.-B. Son (Ed.), *Computer-assisted language learning: Concepts, contexts and practices* (pp. 145-164). APACALL Book: iUniverse, Inc.
- Zhang, F., & Wagner, M. (2005). *Effects of F0 feedback on the learning of Chinese tones by native speakers of English*. Paper presented at the Conference proceeding for Interspeech 2005: Eurospeech-9th European Conference on Speech Communication and Technology, Lisboa, Portugal.
- Zielinski, B. (2006, April 1). The intelligibility cocktail: An interaction between speaker and listener ingredients. *Prospect*, 21, 22-45.