AN EXPLORATORY CASE STUDY TO INVESTIGATE PERCEIVED PRONUNCIATION ERRORS IN THAI PRIMARY SCHOOL STUDENTS USING AUDIO-VISUAL SPEECH RECOGNITION

by **Steven Graham**

American University of Phnom Penh

#278H Street 210R, Kroalkor Village

Sangkat Kilometer 6, Khan Russey Keo

Phnom Penh, Cambodia

shed_chelsea @ hotmail.com

Abstract

An explorative case study has been conducted at a small rural school in the north east of Thailand to investigate the pronunciation errors that primary school students make when reading English aloud. This paper illustrates the opportunities and challenges of employing speech recognition software in rural classrooms by using it with specifically designed audio-visual materials based on the Thai curriculum to identify English language reading and pronunciation difficulties. A comparison is made between this study and published literature.

Keywords: Speech recognition software; audio visual; English; computer assisted language learning

1. Introduction

As researchers understand the brain mechanisms in early language acquisition (Kuhl, 2010), the learning process itself as an infant, not just a period of time spent learning, affects an individual's ability to learn (Kuhl and Rivera-Gaxiola, 2008). In Thailand children learn English from the first grade. There are many theories as to the optimal age when a student should start to learn another language (Kennedy, 2006); however, government legislation dictates that students will learn English from Prathom 1 (Grade 1) in line with research which shows the benefits to a child's development of being bilingual (Barac and Bialystok, 2011). Researchers have discovered how multiple languages are stored in the memory (Riehl, 2010) and that bilinguals have cognitive (Spark, 2010) and job-related advantages over their

counterparts (S. D. Krashen, personal communication, October 1, 2010), so for the time being at least, English is taught from the first grade in Thailand.

English plays an important part in the development of individuals, communities and countries (Coleman, 2010) even though there are some academics that are worried by a perceived threat from Chinese, as Mandarin becomes more popular (Graddol, 2010). Closer to home, Techavijit (2010) advocates "setting" rather than "streaming" in his vision of the direction that education reform should take in the second decade (2009-2018) in Thailand. If this were to take place in conjunction with the sequencing of new critical thinking tasks (Beaumont, 2010) introduced in Curriculum 51 (The Ministry of Education, Thailand, 2008), then there could be a possible pathway out of our present educational demise. The problem is that after nearly 70 years since the first study into critical thinking took place, it is still a long way from being realised in some universities (Elder, 2010), let alone in the first grade.

Shenk (2010) explains that practice may not make perfect; however, ten thousand hours (three hours a day) will help you on your way to "raising the bar" and being successful at what you do. Snyder (1971) informed us forty years ago that we all know that those who learn most are those that use what they have learned. It is with this in mind that we as teachers and learners need to appreciate that learning is a slow and incremental process that can be achieved by anyone given the right opportunity. In order for this to happen, Thailand's schooling must move from a teacher-centred to a more learner-centred approach in line with its 1999 Education Act (Office of the National Education Commission, 2012) and Curriculum 51 (The Ministry of Education, Thailand, 2008) so that our students have a more realistic chance of success. By coaching them and allowing a certain amount of learner autonomy in what they do (Barber and Foord, 2010), students taking part in this study have been given the opportunity to practice speaking English on their own using videos and SpeaKIT voice recognition software which follow the Thai foreign language curriculum.

It is important to understand that the majority of Thailand's primary school teachers of English have majored in subjects other than English and have not had the required training to teach English in accordance with the directives of the 1999 Education Act (Office of the National Education Commission, 2012) and the subsequent Curriculum 51 (The Ministry of Education, Thailand, 2008) documentation. This has resulted in teachers in rural areas admitting continuing to teach in a teacher-centred way as they lack confidence in their own English language skills as well as the techniques needed to pass on knowledge in a student-centred way (Mackenzie, 2002). It was this dilemma that encouraged the development of a series of DVDs called *Smooth Transitions* as a stepping stone to communicative student-

centered English language teaching in Thailand's rural north east and the eventual synthesis with SpeaKIT voice recognition software.

This explorative study asks one question: To what extent does the use of audio-visual communicative activities and speech recognition software in Thailand's rural primary EFL classrooms identify and ameliorate problem areas in students' English language reading and pronunciation?

2. Background

2.1. *Smooth Transitions* DVDs

Graham (2007) proposed a project to train approximately 4,000 Thai teachers of English in a year by using a teacher training framework based on commercial pyramid selling techniques. Several years later, it became evident through constant research and feedback that if any educational project is to be accepted by Thailand's Ministry of Education, a smaller project would need to be piloted first and then undergo extensive scrutiny if it was to have any chance of being adopted on a larger scale.

In order for this to happen, efforts were concentrated on one school (Graham, 2009) in Ban Phue district in Udon Thani province. Bantatprachanukoon School was chosen as the school director was forward thinking and the English language coordinator at that time was very motivated and possessed outstanding English language skills. Together, all parties collaborated successfully resulting in a teacher training project that started with a CD and some flashcards and transformed into the development of six years of basic education learning and teaching materials based on the Thai curriculum and the implementation of SpeaKIT software at the school's ERIC (English Resources Instruction Center) center (Graham, 2010).

It is important to listen to what everyone has to say when there is a collaborative project such as this. Regular feedback sessions (Graham, 2013) motivated the involved teachers. As a result, one of them requested that DVDs be produced with English captions, so that the students had something to watch as well as listen to when they were conducting their dialogues. The central characters in the DVDs are two children dressed in their school uniforms which students can easily identify with. There is also the increased opportunity for students to improve their speech perception as they grow accustomed to the "known voices" (Rost and McMurray, 2010) and the context that the characters are in (Drager, 2010) throughout the series.

Audio visual equipment can be used in many ways. For the initial teacher training project at Bantatprachanukoon School as demonstrated in this study, there was a conscious

effort not to interfere too much with a teacher's existing teaching pedagogy; however, a suggested process was put forward for those teachers who were interested in using it after they had completed their usual method of instruction (Graham, 2010). Teachers would play one section of the DVD as many times as they wanted to so that the students could watch and listen and at the same time the teacher would pause the DVD in order to explain what was happening and confirm that the students understood what was taking place.

Once the teacher believed that the class was ready, they would have the boys repeat the dialogue of the boy on the DVD (Bank) and the girls would follow the girl (Noi). In addition, teachers would then have the students come to the front of the class in pairs (one boy and one girl) and have those students repeat dialogues from the DVD as the teachers paused the DVD at the end of small lexical chunks, ensuring that all students in the class took part in the activity.

Vocabulary substitution is encouraged and highlighted in the written dialogues accompanying the DVDs and within the subtitles of the DVDs themselves allowing teachers the opportunity to use explicit vocabulary instruction to encourage more vocabulary learning after incidental exposure to the DVDs which focuses the learners' attention without straying too far from the specific dialogue (File and Adams, 2010). In addition, further opportunity to use the language closes the gap between what a student knows and what they can say (Cimons, 2010).

For this current study, all participants used pages 3, 17 and 26 from Smooth Transitions DVDs, Prathom 1 for the assessments (see Appendix 1).

2.2. *SpeaKIT* speech recognition software

Cawkell (1999) documents the progress of speech recognition software up until 1999. For those not familiar with this type of technology, it is extremely complex and its intricate workings are outside the scope of this paper; however, some specifications will be explained briefly. The algorithms used are data-driven (Marchand, Adsett and Damper, 2009) and state of the art, with a large vocabulary and a continuous speech recognition system (Lotto and Holt, 2010), more than capable of dealing with the different varieties of English (Franco, Bratt, Rossier, Gadde, Shriberg, Abrash et al., 2010) negating potential problems that can occur concerning the judgements of oral performance assessments conducted by native and non-native teachers (Kim, 2009).

SpeaKIT is beneficial for the project as it has independent automatic speech recognition (ASR) software as opposed to dependent (see Li and Topolewski, 2002, describing their process of designing language learning simulation). Independent ASR means that if the

speaker's speech deviates from the "norm" (that which is deemed acceptable) then it is rejected. In the case of *SpeaKIT*, the user has up to four attempts at self correction before the software moves on to the next word or phrase. Dependant ASR adapts to the user and their pronunciation which would not benefit the user in the same way (Pavlichev, 2002).

The level of proficiency of an English language student has traditionally always been compared to that of native speakers (Muñoz and Singleton, 2011). Amongst others, Richards (2006) points out that today this is not necessarily the case. To this end, *SpeaKIT* has various acceptance levels which were adjusted over time to reflect the required standard for the research project (Li and Topolewski, 2002). Initially, the level was at a low setting to encourage the whole school to use the software and gradually it was increased as the confidence and expectations of success (Dörnyei, 2001) of those using the equipment grew. Interestingly, Ginther, Dimova and Yang (2010) detail the reasons why automated systems are superior to humans at assessment as they possess greater internal consistency.

As in all projects of this kind, it takes a considerable amount of time at the beginning to set up the administration and management functions before it is possible to begin the actual research. Students' names and passwords needed to be input as well as those teachers who acted as managers and administrators. In addition, when students are using the *SpeaKIT* software, only ten active users can be logged in at any one time, causing considerable work for administrators when collecting data in the course of a normal classroom lesson.

Those using the *SpeaKIT* software used the practice, assessment or listening modes to complete their tasks. In the practice mode, students have the opportunity to listen to and repeat or use an advanced mode which allows the student to read and *SpeaKIT* will only interrupt if there is a mistake. As they are speaking, students see bars on the screen indicating whether they have spoken within the allowed boundaries (illustrated by green bars for correct and red for incorrect). They also see a report on the screen when they have finished showing where they have made errors and whether they were able to correct them themselves. The assessment mode allows students to read without interruption even if they make mistakes, with a report generated for the teacher when the activity has finished. The listening mode allows the student just to listen to the text being read. The report on the screen is particularly useful as it provides the teacher with an opportunity to give instant feedback to the students after they have completed the activity.

As *SpeaKIT* is content neutral, teachers are able to write their own content, i.e. lesson material. This can be completed by hand or by importing text or lesson material from elsewhere, which is normally the preferred method. Teachers are also able to use cut and paste

functionality. The software will highlight heteronyms and unknown words allowing the teacher to assign an IPA (International Phonetic Alphabet) designation from the dictionary that is contained in the software and also create unknown words using the IPA and appending these to the dictionary.

3. The study

A collaboration with Bantatprachanukoon School and Udon Thani Education Service Area Office 4 was initiated to integrate dialogues from *Smooth Transitions* with *SpeaKIT*'s speech recognition software. Research was required to investigate how speech recognition software could be used to increase the English language skills of primary school English language learners by identifying errors in reading and pronunciation. Teachers, administrators and managers connected to the project were also included in the planning and implementation process.

A server was set up with four client computers so that there is no need for the internet whilst students were using the computers. Even though these students were from a rural background where the majority would not have access to computers at home, there appeared to be no noticeable demonstration of anxiety (Mcinerney, Marsh and Mcinerney, 1999) since students had previously used computers at the school. Thus, the subject of anxiety was deemed outside the scope of this research and the main project.

3.1. Participants

Thirty-six students from Grades 1-9 (Prathom 1-6 and Mattayom 1-3) at Bantatprachanukoon School were selected as the sample group. The breakdown by classes is shown in Table 1.

Table 1. Sample by class

P1	P2	P3	P4	P5	P6	M1	M2	M3
n=3	n=5	n=5	n=5	n=5	n=5	n=3	n=2	n=3

The class teachers selected some high achievers, some low achievers and some ranked in the middle from the same socio-economic background, so as to attempt to give as true a representation of the makeup of a Thai classroom as possible (Brown, 1988).

3.2. Procedure

In order to maximise the potential of both *Smooth Transitions* and *SpeaKIT*, students taking part in this project at Bantatprachanukoon School continued to use *Smooth Transitions* in the classroom with their teacher over the school year so that they would have already been exposed to the dialogues and have an understanding of what the dialogues meant, the turn taking patterns required, intonation, facial expressions and body posture while speaking English (Goodwin, 2000).

To conform to the process of this investigation, all the students were then required to play the video on the computer as many times as they wanted before switching to the *SpeaKIT* software by using the Alt/Tab keys. Those students who comprised the sample group were required to start at Prathom 1 and to work their way through as many pages or years as they were able to. Students outside the sample group were required to use the Prathom year they were in or lower if it was too difficult for them to complete.

The sample group completed assessments using the existing *SpeaKIT* software in assessment mode which involved students reading three pre-selected pages from the Smooth Transitions DVDs, Prathom 1 (see Appendix 1). The assessment was standardised across the grades in order to be fair due to the low English language levels of the students taking part (Nunan, 1992). A second assessment was given three weeks later. This period of time was too short to be used for pre- and post-test purposes, so the data was collected and used as a whole.

The analysis of the errors was conducted using observations by Smythe (1987), who detailed the phonology problems experienced by Thai people when they try to speak English. His account predominantly concerns L1 interference and details how Thai people stress the final syllables of words, how they have difficulty in pronouncing certain final consonants and final consonant clusters, as well as speaking in a way that gives their English pronunciation a certain staccato effect.

4. Results and discussion

The speech recognition software was able to produce individual reports for each of the thirty-six students of the sample group. Two assessments were administered, three weeks apart and the information was collated in Table 2 to show which words were not pronounced correctly over the two assessments by the sample group.

The first number in brackets is the frequency that the word occurred in the three selected *Smooth Transitions* DVD pages and the second number is the frequency of errors committed by the sample group for that word. The last number is a classification of 100 most frequently used words (Fry, Kress, & Fountoukidis, 2000). The 100 most frequently used words

are split into four groups of twenty-five, where 1 designates the first twenty five and 4 the last. Category 5 is outside the 100 most frequently used words.

Table 2. Total errors made by students from the two assessments

1-5 Errors by Sample	**6-10 Errors by Sample**	**11-14 Errors by Sample**
hello (2, one, 5) /hələʊ/	those (1, six, 5) /ðəʊz/	byebye (1, eleven, 5) /bɑɪbɑɪ/
name (3, one, 5) /neɪm/	too (2, six, 5) /tu:/	chair (1, eleven, 5) /ʧeə/
morning (2, two, 5) /mɔ:nɪŋ/	good (2, seven, 5) /gʊd/	dolls (1, eleven, 5) /dɒlz/
Noi (4, two, 5) /nɔɪ/	nice (1, seven, 5) /nɑɪs/	flower (1, eleven, 5) /flɑʊə/
does (1, three, 2) /dʌz/	apple (1, nine, 5) /æpəl/	father (1, twelve, 5) /fɑ:ðə/
cat (1, four, 5) /kæt/	sofa (1, nine, 5) /səʊfə/	goodbye (3, twelve, 5) /gʊdbɑɪ/
fish (1, four, 5) /fiʃ/	somtam (2, nine, 5) /sɒmtæm/	monkey (1, twelve, 5) /mʌŋki/
mat (1, four, 5) /mæt/	above (1, ten, 5) /əbʌv/	table (1, twelve, 5) /teɪbəl/
meet (2, four, 5) /mi:t/	birds (1, ten, 5) /bɜ:dz/	tree (1, twelve, 5) /tri:/
15-20 Errors by Sample	**21-24 Errors by Sample**	**25+ Errors by Sample**
doll (1, fourteen, 5) /dɒl/	Bank (5, twenty-two, 5) /bæŋk/	vase (1, twenty-five, 5) /vɑ:z/
box (1, fifteen, 5) /bɒks/		
cars (1, fifteen, 5) /kɑ:z/		
pleased (1, seventeen, 5) pli:zd/		
rubber (1,eighteen, 5) /rʌbə/		
likes (1,nineteen, 3) /lɑɪks/		

When looking at vowel sounds in particular, there were twelve words listed as errors that contained sounds that had been classified by Smythe (1987) as not having equivalent sounds in Thai and, thus, they posed a potential problem for Thai speakers of English. As can be seen in Table 3, the first five columns are diphthongs and the last is a triphthong.

Table 3. Word errors listed under phonetic vowel sounds

/eɪ/	/eə/	/ɑɪ/	/əʊ/	/ɔɪ/	/ɑʊə/
name table	chair	nice byebye goodbye likes	hello those sofa	Noi	flower

There were fourteen words listed as errors that contained five consonant sounds that Smythe (1987) considered problematic for Thai speakers of English. In Table 4, they are listed under their respective phonetic sounds.

Table 4. Word errors listed under phonetic consonant sounds

/z/	/ʃ/	/ð/	/g/	/v/
does those birds dolls cars pleased vase	fish	those father	good goodbye	above vase

In addition, consonant clusters at the end of words like *pleased* would also cause difficulties for Thais. This would also include *box* and *Bank*. Furthermore, the backward /l/ sounds of *apple* and *doll* would also prove difficult as well as the consonant cluster at the beginning of the word *tree*. Moreover, *mat* and *meet* can easily be confused due to the length of the vowel sound and *too* has a tendency to have the vowel sound shortened. The word *monkey* has a consonant cluster in the middle of the word and *rubber* tends to have too much stress on the second syllable, as does the word *monkey*. *Somtam* is a Thai word and the author believes that the way the word was given its phonetic transcription was wrong due to the stress put on the second syllable by Thai speakers. The name *Bank* in this dialogue when spoken by Thai speakers of English does not have the final consonant /k/ as a plosive resulting in such a high frequency. The long vowel sounds /ɔ:/ in *morning* and /u:/ in *too* have also historically caused Thais difficulty.

Special consideration has to be given to the word *vase* in this pilot project. The results show twenty five errors; however, on investigation it was found that the speech recognition software has an American pronunciation /veɪs/ and the *Smooth Transitions* DVDs has British /vɑ:z/. The students copied the DVD segment and produced a British pronunciation which was classified as an error by the software.

The words *likes* and *does* are the only words that have been shown as errors from the 100 frequently used words list (Fry, Kress, & Fountoukidis, 2000) used by the *SpeaKIT* program. This is considered somewhat unusual and is the subject of further investigation.

Through exposure to this form of computer supported learning which was in addition to the *Smooth Transitions* DVD classroom activities mentioned earlier, students would improve their higher order thinking skills, social interaction, critical reflection practices and creativity (Ma and Pendergast, 2010), whilst at the same time being allowed more opportunity to read and speak English than they would normally have in their classroom. Feedback at the end of every *SpeaKIT* activity ensured that students had the opportunity to repeat erroneous words correctly with the teacher in a timely manner.

This investigation satisfied the opportunities for second language learning conditions identified by Spolsky (1989), in that it gives the students more reading and speaking time than would be allowed under normal circumstances in the traditional classroom setting. The students who have comprised the sample population demonstrated a particularly good work ethic, as have the coordinating teachers, both internally and externally (Fox and Grams, 2007). This allowed this pilot project to continue even though there were extensive outside work commitments for all cooperating teachers and administrators (Stephenson, 1994).

There is some concern as to only two of the 100 frequently used words (*likes* and *does*) were identified as errors in this pilot project, which will be looked at by the software developers. Notwithstanding this, *SpeaKIT* has already demonstrated, though subject to limitations, that it has the ability to easily adapt any material or curriculum and be able to give students the opportunity to practise and improve their reading and pronunciation skills in a user friendly environment; whilst at the same time linking speech perception and production to acquire new vocabulary (Casserly and Pisoni, 2010). Apart from offering instant feedback on words that were pronounced incorrectly, teachers were also able to incorporate remedial training into the mainstream lessons to overcome the students' language difficulties.

Potcharapanpong and Thongthew (2010) suggest that future teacher training programme developers may consider a "multi-media toolkit or supported-kit" for Thai teachers of English. The use of *SpeaKIT* with the *Smooth Transitions* DVDs could go some way to reinforcing their idea, as teachers are in a position to benefit from the use of these materials in the same ways as their students. In addition, sociopragmatic competence is further enhanced by having the opportunity to see the body language that accompanies the various politeness strategies in the newly acquired L2 on the *Smooth Transitions* DVDs before having the opportunity to read it and speak it using the speech recognition software (Phillips, 1993).

As Aziz (2010) explains, there are five keys to educational technology, considered implementation of appropriate tools, techniques, or processes that facilitate the application of senses, memory, and cognition to enhance teaching practices and improve learning outcomes. Technology assists teachers in the classroom rather than replace them and has already proved to be a great motivator for both teachers and students (Quinn, 2007). By integrating *Smooth Transitions* and *SpeaKIT* with an existing course book that follows the Thai foreign languages curriculum, students and teachers have the opportunity to concentrate on all four language skills, implicitly and explicitly, with an increased focus on reading and speaking depending on their needs, so they become "capable, independent learners in an efficient manner" (Shen, 2003).

Conducting a study like this in a rural area involved trying to overcome many problems; one of which is the availability of an internet connection. Many schools have an internet facility; however, this service at Bantatprachanukoon School seemed to be out of order frequently and for long periods of time which complicated some of the management and administrative tasks required to make the project work. Thailand is not alone in this problem. Nigeria experiences similar difficulties (Dala, 2009); however, with proper funding, well maintained equipment and dedicated staff, this project managed to function giving the children, teachers and staff of Bantatprachanukoon School the opportunity to practise their English in an autonomous and relaxed environment (Thanasoulas, 2000).

Further research is required into phonetic settings by L2 researchers (Mennen, Scobbie, Leeuw, Schaeffler and Schaeffler, 2010) to enable the academic community to truly understand articulation. The word *somtam* is a point in question as the way a Thai person pronounces the word is considerably different from how it was input into the *SpeaKIT* program. The same can be said for the name *Bank*.

In addition, investigation into the use of analogical mapping models to assist in understanding human cognition and improve education and training (Gentner and Forbus, 2010) would be beneficial as well as further research is needed into the effectiveness of computer assisted response analysis schemes (Chapelle and Chung, 2010) for the purpose of assessment.

5. Conclusion

As Pinter (2006) reminds us, Piaget (1923) explained how children of a similar age have similar characteristics. Vygotsky (1978) stated that social interaction between children and their parents and teachers offer unique valuable experiences. If we add Gardner (1993) and multiple

intelligences to the equation, it is apparent that audio visual speech recognition has the potential to have something for every child and that it is up to the teacher to identify the student's individual learning style and adapt the implementation of the program according to the needs of the students in their specific classroom. To this end, future studies should pay particular attention to learner preferences and conduct interviews and surveys with teachers and students to see exactly what students want to learn and how they want to learn it.

There is great potential for teachers to use the audio-visual speech recognition for their own English language skills development as well as for lesson preparation. This is another area for practise and future research.

This project has demonstrated that the integration of the curriculum-based *Smooth Transitions* and the openness of *SpeaKIT* allows schools to create content to identify reading and pronunciation errors, i.e. lesson material based on the national curriculum, whilst also addressing local cultural needs. What is required now is detailed research into the benefits of integrating video with speech recognition software (audio visual speech recognition) into Thailand's primary English language classrooms and to identify any additional requirements that may be needed by those using the facility.

Acknowledgements

The author would like to thank Channarong Rachbuanoy from the Udon Thani Primary Educational Service Area Office 4 for his assistance with administration and Prudchayaphorn Boontool for her data collection, attention to detail, and patience. In addition, Udon Education Foundation supplied the computer hardware for this project and Princeton Language School the technical support for the *SpeaKIT* software.

Funding

This research was supported by a scholarship from Khon Kaen University, Khon Kaen, Thailand.

References

Aziz, H. (2010). *The 5 keys to educational technology*. Retrieved from http://thejournal.com/articles/2010/09/16/the-5-keys-to-educational-technology.aspx

Barac, R., & Bialystok, E. (2011). Cognitive development of bilingual children. *Language Teaching, 44*(1), 36-54. DOI: 10.1017/S0261444810000339.

Barber, D., & Foord, D. (2010). Learning coach. *English Teaching Professional*, 71, 53-54.

Beaumont, J. (2010). A sequence of critical thinking tasks. *TESOL Journal, 1*(4), 427-448. DOI: 10.5054/tj2010.234763.

Brown, J. D. (1988). *Understanding research in second language learning*. Cambridge: Cambridge University Press.

Casserly, E. D., & Pisoni, D. B. (2010). Speech perception and production. *WIREs Cognitive Science, 1*(5), 629-647. DOI: 10.1002/wcs.63.

Cawkell, T. (1999). Tracking fast-moving technology: The progress of speech recognition. *Journal of Information Science, 25*(1), 79-84. DOI: 10.1177/016555159902500108.

Chapelle, C. A., & Chung, Y-R. (2010). The promise of NLP and speech processing technologies in language assessment. *Language Testing, 27*(3), 301-315. DOI: 10.1177/0265532210364405.

Cimons, M. (2010). *Understanding language development*. Retrieved from http://www.usnews.com/science/articles/2010/07/22/understanding-language-development?PageNr=1

Coleman, H. (2010). *The English language in development*. Retrieved from http://www.teachingenglish.org.uk/sites/teacheng/files/UK011-English-Language-Development.pdf

Dala, H. S. (2009). The role of information and communication technology in academic libraries and the achievement of the millennium development goals (MDG). *The Voice of Teachers, 1*(2), 141-144.

Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.

Drager, K. (2010). Sociophonetic variation in speech perception. *Language and Linguistics Compass 4(7)*. DOI: 10.1111/j.1749-818X.2010.00210.x. Retrieved from http://www.blackwell-compass.com/subject/linguistics/article_view?parent=section&last_results=section%3Dlnco-sociolinguistics&sortby=date&section=lnco-sociolinguistics&browse_id=lnco_articles_bpl210&article_id=lnco_articles_bpl210

Elder, L. (2010). *Achieving critical mass*. Retrieved from http://www.timeshighereducation.co.uk/story.asp?sectioncode=26&storycode=414351

File, K. A., & Adams, R. (2010). Should vocabulary instruction be integrated or isolated? *TESOL Quarterly, 44*(2), 222-249. DOI: 10.5054/tq.2010.219943.

Fox, W. S., & Grams, C. L. (2007). Work ethic as school-based behaviours of adolescent students. *Family & Consumer Sciences Research Journal, 36*(1), 63-87. DOI: 10.1177/1077727X07303819.

Franco, H., Bratt, H., Rossier, R., Gadde, V. R., Shriberg, E., Abrash, V., et al. (2010). EduSpeak: A speech recognition and pronunciation scoring toolkit for computer-aided language learning applications. *Language Testing, 27*(3), 401-418. DOI: 10.1177/0265532210364408.

Fry, E. B., Kress, J. E., & Fountoukidis, D. L. (2000). *Reading teacher's book of lists.* San Franscisco: Jossey-Bass.

Gardner, H. (1993). *Multiple intelligences new horizons*. New York: Basic Books.

Gentner, D., & Forbus, K. D. (2010). Computational models of analogy. *WIREs Cognitive Science, 2*(3), 266-276. DOI: 10.1002/wcs.105.

Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing, 27*(3), 379-399. DOI: 10.1177/0265532210364407.

Goodwin, C. (2000). Action and embodiment within human situated interaction. *Journal of Pragmatics, 32*(10), 1489-1522. DOI: 10.1016/S0378-2166(99)00096-X.

Graddol, D. (2010). Will Chinese take over from English as the world's most important language? *English Today, 26*(4), 3-4.

Graham, S. (2007). *Proposed teacher training for primary school teachers of English in Thailand*. Retrieved from http://www.steves-english-zone.com/images/Papers/proposed-teacher-training-for-primary-school-teachers-of-english-in-thailand.pdf

Graham, S. (2009). From the bottom up: A case study of teacher training for primary school teachers of English in a Thai school in north eastern Thailand. *English Language Teacher Education and Development, 12*, 31-43.

Graham, S. (2010). EIL in the primary classroom: Exploration and innovation using DVDs for communication. *Journal of Asia Pacific Studies, 1*(3), 506-523.

Graham, S. (2013). Reflection and feedback: Primary school teachers of English take control. *Humanising Language Teaching 15*(5). Retrieved from http://www.hltmag.co.uk/oct13/mart01.htm

Kennedy, T. J. (2006). Language learning and its impact on the brain: Connecting language learning with the mind through content-based instruction. *Foreign Language Annals, 39*(3), 471-486. DOI: 10.1111/j.1944-9720.2006.tb02900.x.

Kim, Y-H. (2009). An investigation into native and non-native teachers' judgments of oral English performance: A mixed methods approach. *Language Testing, 26*(2), 187-217. DOI: 10.1177/0265532208101010.

Kuhl, P., & Rivera-Gaxiola, M. (2008). Neural substrates of language acquisition. *The Annual Review of Neuroscience, 31*, 511-534. DOI: 10.1146/annurev.neuro.30.051606.094321.

Kuhl, P. K. (2010). Brain mechanisms in early language acquisition. *Neuron, 67*(5), 713-727. DOI: 10.1016/j.neuron.2010.08.038.

Li, R-C., & Topolewski, D. (2002). Zip & Terry: A new attempt at designing language learning simulations. *Simulation & Gaming, 33*(2), 181-186. DOI: 10.1177/1046878102332006.

Lotto, A., & Holt, L. (2010). Psychology of auditory perception. *WIREs Cognitive Science,* 2(5), 479-489. DOI: 10.1002/wcs.123.

Ma, A., & Pendergast, D. (2010). Innovative pedagogies for family and consumer science/home economics education - Utilising computer-based collaborative learning to foster lifelong learning attributes. *Family & Consumer Sciences Research Journal, 38*(3), 273-288. DOI: 10.1111/j.1552-3934.2009.00018.x.

Mackenzie, A. S. (2002). EFL curriculum reform in Thailand. Paper presented at the Curriculum Innovation, Testing and Evaluation: Proceedings of the 1st Annual KALT Pan-Sig Conference, Kyoto, Japan.

Marchand, Y., Adsett, C. R., & Damper, R. I. (2009). Automatic syllabification in English: A comparison of different algorithms. *Language and Speech, 52*(1), 1-27. DOI: 10.1177/0023830908099881.

Mcinerney, V., Marsh, H. W., & Mcinerney, D. M. (1999). The designing of the computer anxiety and learning measure (Calm): Validating of scores on a multidimensional measure of anxiety and cognitions relating to adult learning of computing skills using structural equation modeling. *Educational and Psychological Measurement, 59*(3), 451-470. DOI: 10.1177/00131649921969974.

Mennen, I., Scobbie, J. M., de Leeuw, E., Shaeffler, S., & Schaeffler, F. (2010). Measuring language-specific phonetic settings. *Second Language Research, 26*(1), 13-41. DOI: 10.1177/0267658309337617.

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching, 44*(1), 1-35. DOI: 10.1017/S0261444810000327.

Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.

Office of the National Education Commission. (2012). *National Education Act 1999.* Retrieved April 12, 2015, from Office for National Education Standards and Quality Assessment: http://www.onesqa.or.th/en/publication/nation_edbook.pdf

Pavlichev, A. (2002). Software review. *Social Science Computer Review, 20*(1), 93-98. DOI: 10.1177/089443930202000109.

Phillips, E. M. (1993). Polite requests: Second language textbooks and learners of French. *Foreign Language Annals, 26*(3), 372-381. DOI: 10.1111/j.1944-9720.1993.tb02293.x.

Piaget, J. (1923). *The language and the thought of the child*. New York: Harcourt Brace and World.

Pinter, A. (2006). *Teaching young language learners*. Oxford: Oxford University Press.

Potcharapanpong, S., & Thongthew, S. (2010). TEFL training program for local and cultural-based instruction: Case studies in Thai primary schools. *Asian Social Science, 6*(9), 39-48.

Quinn, Q. (2007). *Technology and motivation: Can computers motivate students to read?* Retrieved from http://www.compasslearning.com/CompassFileUpload/60Technology_Motivation.pdf

Richards, J. C. (2006). Materials development and research—making the connection. *RELC Journal*, *37*(1), 5-26. DOI: 10.1177/0033688206063470.

Riehl, C. M. (2010). The mental representation of bilingualism. *WIREs Cognitive Science, 1*(5), 750-758. DOI: 10.1002/wcs.74.

Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy, 15*(6), 608-635. DOI: 10.1111/j.1532-7078.2010.00033.x.

Shen, H-J. (2003). The role of explicit instruction in ESL/EFL reading. *Foreign Language Annals, 36*(3), 424-433. DOI: 10.1111/j.1944-9720.2003.tb02124.x.

Shenk, D. (2010). *The genius in all of us*. New York: Doubleday.

Smythe, D. (1987). Thai speakers. In M. Swan, & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (pp. 252-263). Cambridge: Cambridge University Press.

Snyder, J. W. (1971). Who learns most? *Foreign Language Annals, 4*(4), 379-384. DOI: 10.1111/j.1944-9720.1971.tb02054.x.

Spark, S. (2010). *Science grows on acquiring new language.* Retrieved from http://www.edweek.org/ew/articles/2010/10/22/09window_ep.h30.html?tkn=NSOF+.com

Spolsky, B. (1989). *Conditions for second language learning*. Oxford: Oxford University Press.

Stephenson, H. (1994). Management and participation in ELT projects. *ELT Journal, 48*(3), 225-232. DOI: 10.1093/elt/48.3.225.

Techavijit, V. (2010, December 10). Only happy schools produce well rounded leaders. *Bangkok Post.* Retrieved from http://www.bangkokpost.com/learning/education-features/210701/only-happy-schools-produce-well-rounded-leaders

Thanasoulas, D. (2000). What is learner autonomy and how can it be fostered? *The Internet TESL Journal, 6(11)*. Retrieved from http://iteslj.org/Articles/Thanasoulas-Autonomy.html

The Ministry of Education, Thailand. (2008, July 11). *Assumption College Thonburi.* Retrieved March 5, 2017, from http://www.act.ac.th/document/1741.pdf

Vygotsky, L. (1978). *Mind and society, the development of higher mental processes*. Cambridge: Harvard University Press.

Appendix 1. Pre-selected pages from *Smooth Transitions*

Prathom 1 Page 3

Good morning. My name is Bank. What is your name?

Good morning. My name is Noi. Nice to meet you.

Pleased to meet you too. What is this?

This is a rubber. What is that?

That is a chair. Goodbye Noi.

Goodbye Bank.

Prathom 1 Page 17

Hello Bank.

Hello Noi. What is it?

It is an apple. What are those?

They are fish. What are these?

They are dolls. What do you like?

I like somtam. What do you like?

I like somtam too. What does your father like?

He likes cars. Goodbye Noi.

Bye-bye Bank.

Prathom 1 Page 26

The monkey is on the table.

The cat is in the box.

The doll is on the sofa.

The flower is in the vase.

Bank is on the mat.

The birds are above the tree.