The Use of Web-Based Text-to-Speech Tool in Improving English Pronunciation and Changing Perception of Error Correction among Young Learners

Ian James

SK Tombovo, Penampang, Sabah, Malaysia
ianflojam@yahoo.com

ABSTRACT

Text-to-speech is a web-based tool that has the potential to improve English as a Second Language (ESL) learners' spoken production. This research explored the role of that tool in helping reluctant speakers to improve their pronunciation. It also explored its role in changing their perception of error correction. The research was conducted in a Malaysian primary school in a small village. It was targeted at 9 children who had been identified as proficient writers but reluctant speakers. In the first cycle of the research, the text-to-speech tool was used as a method of instruction. It was used to assist them in pronouncing words correctly by producing conversation scripts and converting them into spoken output through a web-based text-to-speech tool (http://www.ivona.com). In the second cycle, the same activity was carried out but a more playful approach was applied. Throughout these two cycles, the respondents' pronunciation and perception of error correction were observed and evaluated. The data in this research were gathered through audio recording, transcription, observation, interviews, and journal entries. The findings of this research reveal that the use of a web-based text-to-speech tool can improve children's pronunciation because it enables them to use authentic and conversational language. Secondly, the use of this tool can improve children's perception of error correction by means of a non-threatening, self-regulated learning environment. This research also reveals that the use of a web-based text-to-speech tool can be developed if it is combined with various strategies such as gamification, positive reinforcements, and cooperative learning.

Keywords: text-to-speech, pronunciation, error correction, young learners

INTRODUCTION

According to MacCarthy (1998) and Nunan (2001), mastering speaking abilities is the ultimate goal of acquiring a foreign or second language and the other skills are overshadowed by its significance. Along similar lines, Bygate (1987) states that it is by this skill that learners are judged as it is the vehicle par excellence of social solidarity, of social ranking, of professional advancement and of business.

Despite that, speaking has become a neglected skill in Malaysian classrooms. Most ESL teachers would rather focus on reading and writing as their students' performances in these skills are tested in formal examinations such as Ujian Pencapaian Sekolah Rendah (UPSR), Penilaian Menengah Rendah (PMR) and Sijil Pelajaran Malaysia (SPM) (Hassan & Selamat, 2002). The consequence of this is that most Malaysian students become good writers but poor speakers of English. As revealed by The Cambridge Baseline Study (2013), speaking emerged as the weakest skill for Malaysian students at all school grades. Along similar lines, The Malaysian Insider (2004) reported that there existed a big number of job seekers with *A* in SPM English but could not speak a word of it.

Another reason why most ESL teachers avoid speaking lessons is that they do not want to demotivate their students by providing too much error correction. Their view on this matter is consistent with that of Ur (1996) who argues that, while error correction may be valuable to language learning, too much of it can be discouraging and demoralising. Likewise, Parnell (1989), Wadensjö, Dimitrora, and Nilsson (2007), Mishra (2005), and Budden (2009) argue that teachers need to provide error correction at the end of any speaking-related activity so as to avoid undue interruption which might produce a demotivating effect on students.

In view of these issues, it is about time teachers found an appropriate form of intervention to improve their students' oral proficiency and to help develop their students' perception of error correction. Taking into account Barron's (2002)'s view that technology has become an integral and viable part of learning for today's students, the use of a web-based tool can be seen as the appropriate intervention because 'online learning provides flexibility of access to material anywhere anytime, allowing learners to collapse time and space' (Cole, 2000). Moreover, given Buckley and William's (2010) view that 'the use of web technologies provide an opportunity for students to explore their own understanding within a supportive and non-threatening environment', it can be seen as an effective strategy to develop the students' positive perception towards error correction.

On the whole, this action research aimed at examining the strategy that I could employ in my speaking lessons. Having chosen a web-based text-to-speech tool as my intervention, it was my interest to examine the extent to which it had an impact on my students' pronunciation. This research also sought to determine how the web-based text-to-speech tool influenced my students' attitudes towards error correction. Specifically, this research addressed the following three questions:

RQ1: What impact does a web-based text-to-speech tool have on children's speech the segmental and suprasegmental?

RQ2: How does a web-based text-to-speech tool influence children's perception of error correction?

RQ3: How can a web-based text-to-speech tool develop children's pronunciation and change their perception of error correction?

BACKGROUND

This research focused on the respondents' speech accuracy. The British Council (2014) defines accuracy as 'the correct use of the language system, including the use of grammar, pronunciation and vocabulary.' For the purpose of this research, I only focused on the respondents' pronunciation. I chose to focus on pronunciation for a number of reasons. Firstly, it appeared to be my students' most significant weakness when it came to their spoken production. Secondly, as pointed out by Rizvi (2005), good pronunciation is one of the aspects that should be taken into account in order to improve self-expression and achieve the desired clarity and fluency.

For the purpose of dealing with my students' pronunciation, I chose to focus on two aspects, namely segmental (vowels, consonants and diphthongs) and suprasegmentals (word stress and sentence intonation). This is in line with Pascoe, Stackhouse and Wells's (2006) view that both segmental and suprasegmental factors influence one's intelligibility in speech.

It is worth noting though that, on the subject of speech segmental, time constraints made it impossible for me to address all the 24 consonants, 14 vowels and 7 diphthongs in the English language. Hence, the respondents' speech segmental was addressed only when they committed errors in their pronunciation. Likewise, when dealing with their speech suprasegmentals, I only highlighted the elements of word stress and sentence intonation when the respondents produced any type of stress or intonation errors in their speeches.

Given that all the elements of segmental and suprasegmentals were addressed on an 'as-it-happens' basis, error correction is bound to happen frequently over the course of this research. For this reason, it was also my interest to focus on the respondents' perception of error correction. This means that, throughout the process of implementing my intervention, the respondents' reactions when discovering their own mistakes by means of the web-based text-to-speech tool were closely observed and evaluated.

THE STUDY

Target Group

This research involved 9 Year Six students aged 12 years old. Based on their writing ability, the students were identified to have average to high English language proficiency. Hence, the academic performances of the students ranged from good to excellent. Out of the 9 students, 6 were boys and 3 were girls. The parents/guardians of the children involved in this research were in the medium-income group. The highest academic level of the parents/guardians of the respondents were either secondary school or tertiary education.

Research Method

This research was conducted for a duration of two weeks and involved two cycles. The first cycle was conducted to examine the impact of a web-based text-to-speech tool on the respondents' speech segmental and suprasegmentals whereas the second cycle was conducted to examine whether the slightly different approach used had any impact on the respondents' speech segmental and suprasegmentals. Throughout the two cycles, the respondents' perception of error correction and the inherent features of the intervention were also investigated. This research was conducted based on Susman's Model of Action Research which involved 5 stages, i.e. Diagnosing, Action Planning, Taking Action, Evaluating and Specifying Learning. The diagram (Figure 1) below illustrates the structure of implementation for my research.

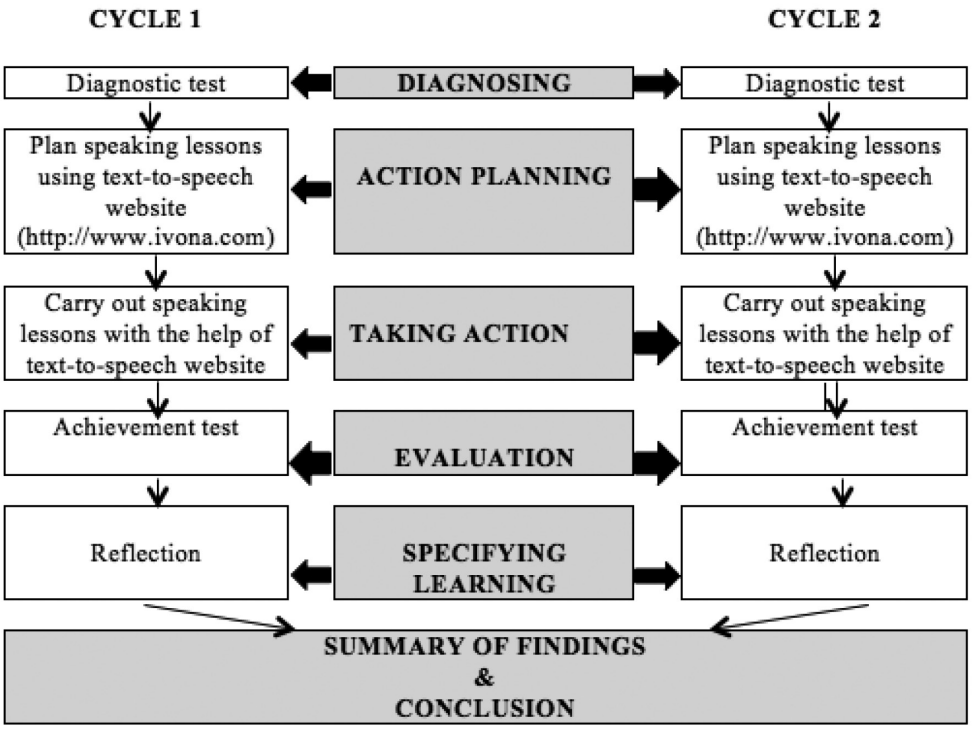

Figure 1 Structure of Implementation for Research (based on Model Susman's, 1983)

Data Collection and Analysis Method

RQ1: *What impact does a web-based text-to-speech tool have on children's speech segmental and suprasegmentals?*

Table 1 Data collection and analysis methods for RQ1

Data collection methods	Data analysis methods
Audio recording	The respondents' speeches were recorded to enable me to gather data related to the first research question. To ensure the validity and reliability of my audio recording, I used a digital audio-recording device which had a much higher signal to noise ratio. In the event that the clarity of the recording was deemed insufficient (e.g. the /θ/ sound could be misheard as the /f/ sound), I got the said respondents to utter the required words and validated his/her pronunciation by paying close attention to the position of his/her tongue, teeth or mouth. To analyse the audio recording, the method that I employed was data reduction. This means that, I simplified the data by only transliterating the respondents' erroneous utterances, i.e. speech patterns that contained segmental and suprasegmentals mistakes.
Transcription	For the purpose of facilitating my analysis of the respondents' speech segmental and suprasegmentals, the recording of the students' erroneous utterances were transcribed. For speech patterns that contained segmental mistakes, the transcription consisted of phonetic alphabets; whereas for speech patterns that contained suprasegmental mistakes, their word stress and intonation were marked. The validity and reliability of my transcriptions were established by making sure that the International Phonetic Alphabet (IPA) was used. Since the IPA is internationally recognised, my transcriptions would provide valuable evidence for my research since the data would be accessible to anyone interested in verifying my findings. Besides, all my evaluations were based on trustworthy, reputable resources, i.e. The Oxford English Dictionary, The Cambridge Advanced Learner's Dictionary, and J.C. Well's (2006) English Intonation. To analyse the respondents' speech transcripts, the method that I employed was codifying. This means that the data were segregated and grouped according to their types of mistakes, i.e. for segmental mistakes – consonants, vowels and diphthongs; while for suprasegmental mistakes – word stress and sentence intonation.
Journal entry	My journal entry consisted of my observations, feelings and insights on each of the respondents' speech segmental and suprasegmentals. To analyse the data from my journal entry, the deductive constant comparison method was employed. This involved chunking my reflective accounts into smaller meaningful parts, and transferring those chunks to a checklist containing a number of criteria and descriptors. The validity and reliability of my journal entry were established by making sure both the checklists for speech segmental and suprasegmentals were adapted from a reputable source, i.e. www.english-grammar-revolution.com. Besides, all the data were collected strictly based on the data gathered earlier from the audio recording and speech transcripts.

RQ2: *How does a web-based text-to-speech tool influence children's perception of error correction?*

Table 2 Data collection and analysis methods for RQ2

Data collection methods	Data analysis methods
Observation	One of the methods I used to gather the data related to Research Question 2 was observation. This means that the respondents' perception towards error correction was monitored closely throughout the implementation of my intervention strategies. For the sake of validity and reliability, the observation was conducted by three people – me and two of my colleagues, both of whom were nicknamed Colleague 1 and Colleague 2. The data from both my and Colleague 1's observation were recorded in written form. Colleague 1's written account was analysed using the inductive analysis method, in that her observation notes were read thoroughly and any significant chunk, phrase or sentence was underlined. Then each of those chunks/phrase/sentences were assigned a code. In the final stage, the codes were combined and the connections between them were identified based on Spradley's (1979) Universal Semantic Relationships.
Journal entry	This was my written account. It consisted of my feelings and insights on the respondents' perception of error correction. To analyse the data from my journal entry, the deductive constant comparison method was employed. This involved chunking my reflective accounts into smaller meaningful parts, and transferring those chunks to a checklist containing a number of criteria and descriptors. The validity and reliability of my journal entry were established by making sure the checklist was adapted from a reputable source, i.e. the International Centre for Leadership in Education's Student Engagement Walkthrough Checklist.
Interview	As previously discussed, only Colleague 1 and my observations were recorded in written form. Colleague 2's observation, on the other hand, was discussed through my interview sessions with her. My interview questions centred on her opinions on the respondents' perception of error correction as well as the observation notes written by Colleague 1. In order to enable me to obtain richer information from Colleague 2, the type of interview that I employed was semi-structured. I believed this type of interview gave me more flexibility as I had the freedom to ask her to explain further when her answer prompted me to learn more about a certain issue. For analysis purposes, our interview sessions were transcribed and analysed using the inductive comparison method as explained above.

RQ3: *How can a web-based text-to-speech tool develop children's pronunciation and change their perception of error correction?*

For the purpose of answering Research Question 3, I chose to apply the same data collection methods as Research Question 2, i.e. observation, journal entry and interview. However, there were two differences. Firstly, all of my, Colleague 1 and Colleague 2's observations not only centred around the respondents' perception of error correction, but also on the inherent features of the web-based text-to-speech

tool. Secondly, all the data from Colleague 1's written account, my journal entry, and the transcripts of my interview with Colleague 2 were analysed using the previously-explained inductive comparison method.

RESULTS AND DISCUSSION

Cycle 1 – Web-based Text-to-Speech as a Method of Instruction

RQ1: *What impact does a web-based text-to-speech tool have on children's speech segmental and suprasegmentals?*

Based on the audio recording, transcription and journal entries' findings in both Cycles 1 and 2 (see Tables 3 and 5), it is found that web-based text-to-speech tool can indeed improve children's word stress and sentence intonation. Thus, it can be concluded that teachers can utilise it if they wish to develop their students' speech suprasegmentals. As for speech segmental, it is found that web-based text-to-speech tool has the potential to improve children's pronunciation of vowel, consonant and diphthong sounds. However, there are several things that teachers should bear in mind prior to implementing it in their speaking lessons.

Firstly, it cannot be carried out as simply a method of instruction. The reason is that it might only develop their students' pronunciation of segments with which they are familiar (e.g. vowel sounds). Besides, they might create a demotivating learning environment which, in due course, would affect their students' willingness to improve themselves.

Secondly, teachers cannot expect their students to improve their pronunciation by merely instructing them to use the web-based text-to-speech tool. It is because there is a high possibility that their students might forget what they have been taught and resort to making the same mistakes again. This, in turn, might cause teachers to give an excessive amount of corrective feedback and thereby demotivating their students.

These findings are consistent with the view put forward by Clark (2003) that the success of a web-based instruction is directly proportional to the success a classroom instruction.

Based on these realisations, teachers should take several measures in order to utilise the web-based text-to-speech tool to its fullest potential.

Creating a motivating learning environment is one of the measures that can be taken. This can be done by gamifying the use of the web-based text-to-speech tool. Kapp (2012) defines gamification as the application of typical elements of game playing like

point scoring, competition with others and rules of play. With reference to the web-based text-to-speech tool, teachers could divide their students into several groups, encourage them to compete with each other, and reward the best-performing group. The data collected from this research have proved that, by doing all these, children are able to improve their pronunciation of segments which they initially find difficult. It is because they would have a stronger desire to perform to the best of their ability. In addition, they are less likely to make the same mistakes when they are highly motivated. This, to a great extent, supports the view put forward by Gordon (2005) that students will learn better when they help teach one another than they will in completely teacher-directed classrooms.

Table 3 The impact of a web-based text-to-speech tool on children's speech segmental and suprasegmental (Cycle 1)

	Audio recording	Transcriptions	Journal entries
Segmental	Web-based text-to-speech tool has a **positive impact** on children's speech segmental	Web-based text-to-speech tool has a **positive impact** on children's pronunciation of **vowel sounds** (a) Web-based text-to-speech tool has a **negative impact** on children's pronunciation of **consonant sounds** (b) Web-based text-to-speech tool has a **mixed impact** on children's pronunciation of **diphthong sounds**	(a) Web-based text-to-speech tool has **a positive impact** on children's pronunciation of **vowel sounds** (b) Web-based text-to-speech tool has **little/ no impact** on children's pronunciation of **consonant sounds** (c) Web-based text-to-speech tool has a **mixed impact** on children's pronunciation of **diphthong sounds** (d) Children make mistakes when they are overexcited
Supraseg-mentals	Web-based text-to-speech tool has a **positive impact** on children's **word stress** and **sentence intonation**	Web-based text-to-speech tool has a **positive impact** on children's **word stress** and **sentence intonation**	Web-based text-to-speech tool has a **positive impact** on children's **word stress** and **sentence intonation**

Table 4 The ways in which web-based text-to-speech tool influences children's perception of error correction

Journal entries	Observation	Interview
Web-based text-to-speech tool contributes much to the positive perception of error correction among children	Web-based text-to-speech tool contributes much to the positive perception of error correction among children	Web-based text-to-speech tool contributes much to the positive perception of error correction among children

Cycle 2 – The Playful Approach of Speaking Lesson with the Help of Web-Based Text-to-Speech Tool

Table 5 The impact of web-based text-to-speech tool on children's speech segmental and suprasegmentals (Cycle 2)

	Audio recording	Transcriptions	Journal entries
Segmental	Web-based text-to-speech tool has a **positive impact** on children's speech segmental	Web-based text-to-speech tool has a **positive impact** on children's speech segmental	Web-based text-to-speech tool has a **positive impact** on children's speech segmental
Supraseg-mentals	Web-based text-to-speech tool has a **positive impact** on children's **word stress** and **sentence intonation**	Web-based text-to-speech tool has a **positive impact** on children's **word stress** and **sentence intonation**	Web-based text-to-speech tool has a **positive impact** on children's **word stress** and **sentence intonation**

Table 6 The way in which Web-based Text-to-Speech Tool can develop children's pronunciation and change their perception of error correction

Journal entries	Observation	Interview
The respondents' improvement in oral proficiency and perception of error correction is a result of the use of web-based text-to-speech tool through motivating activities	Web-based text-to-speech tool can improve children's oral proficiency and perception of error correction because it exposes them to authentic and conversational language in a non-threatening manner	1. Web-based text-to-speech tool can improve children's oral proficiency and perception of error correction because: (a) children are exposed to authentic and conversational language (b) children's confidence in speaking is developed (c) children can enrich their vocabulary through incidental learning (d) it promotes self-monitoring and self-correction 2. The effectiveness of scripted role play can be enhanced by means of motivating activities, cooperative learning, and positive reinforcements

Last but not least, teachers should make a constant effort to give positive reinforcements to their students by approving, encouraging and praising them accordingly. Based on the findings of this research, it is found that children will try harder to improve themselves when their correct utterances are acknowledged and praised. One potential problem that might arise, however, is that teachers might disrupt their students' conversations. Therefore, to avoid undue interruption, teachers can use 'non-

verbal reinforcers' as suggested by McNamara (2014) such as eye contact and friendly expression; standing close to a pupil, nodding while scrutinising work; and a 'thumbs up' sign or other esoteric signals.

RQ2: *How does web-based text-to-speech tool influence children's perception of error correction?*

The data gathered from my journal entries, observation, and interview (see Table 4) suggest that the web-based text-to-speech tool contributes much to the positive perception of error correction. This is because, by means of self-regulated online learning, children can improve themselves through self-correction rather than teacher-correction. This finding is consistent with a previous study conducted by Agudo in 2014. In the study, it was discovered that a high percentage of the participating respondents believed that self-correction would make a greater contribution to reducing their stress and anxiety as opposed to peer-correction and teacher-correction.

Having said that, children's perception of error correction can further be improved if:

(a) their teacher knows how to organise his classroom activities efficiently
(b) their teacher gives them positive reinforcements on a regular basis
(c) they learn in a playful and motivating environment

RQ3: *How can the web-based text-to-speech tool develop children's pronunciation and perception of error correction?*

The findings of this research (see Table 4) reveal that web-based text-to-speech tool can develop children's oral proficiency and error correction because it enables them to use authentic and conversational language confidently by means of a non-threatening learning environment, as well as to enrich their vocabulary through incidental learning.

In order to further improve children's perception of error correction, their teachers must incorporate motivating activities, cooperative learning and positive reinforcements in the lessons.

CONCLUSION

In summary, this research has enabled me to see that a web-based text-to-speech tool indeed has the potential to improve children's oral proficiency because it enables students to use authentic and conversational language in a non-threatening learning environment, as well as enrich their vocabulary through incidental learning. Besides that, children's

perception of error correction can be improved because, by means of self-regulated online learning, children have the ability to correct their own mistakes without the interference of their teachers or peers.

However, when it is executed solely as a method of instruction, teachers might not be able to utilise it to its fullest potential because, in the absence of a motivating learning environment, their students might eventually forget what they have learned. On the other hand, by incorporating the elements of motivation (e.g. gamification and positive reinforcements), and cooperative learning (e.g. competition with others), the effectiveness of a web-based text-to-speech tool can be significantly enhanced and, in addition, children will have a better perception of error correction.

Overall, it can be theorised that: (a) the use of a web-based tool can enhance children's oral proficiency; (b) self-regulated online learning improves children's perception of error correction in that it promotes the practice of self-correction in a non-threatening environment; and (c) the degree of motivation is directly proportional to the quality of learning.

REFERENCES

Agudo, M. (2014). *English as a foreign language teacher education: Current perspectives and challenges*. Amsterdam: Rodopi.

Barron, A. E. (2002). Benefits of technology in education. In A. E. Barron, G. W. Orwig, K. S. Ivers, & N. Lilavois, *Technologies for Education: A practical guide* (p. 3). Santa Barbara: Libraries Unlimited.

British Council. (2014). Teaching English: Retrieved from http://www.teachingenglish.org.uk/knowledge-database/accuracy

Buckley, C. N., & Williams, A. M. (2010). Web 2.0 technologies for problem-based and collaborative learning: A case study. In I. R. USA, *Web-based education: Concepts, methodologies, tools, and applications* (p. 945). Hershey: Idea Group Inc (IGI).

Budden, J. (2009). *Teen world: Multi-level photocopiable activities for teenagers.* Cambridge: Cambridge University Press.

Bygate, M. (1987). Speaking. In C. Carter, & D. Nunan, *The Cambridge guide to teaching English to speakers of other languages.* Cambridge: Cambridge University Press.

Cambridge Advanced Learners Dictionary. (2006). Stuttgart: Klett.

Clark, R. E. (2003). Research on web-based learning. In R.H. Bruning, C. A. Horn, & L. M. PytlikZillig, *Web-based learning: What do we know? Where do we go?* (p.3). Charlotte, North Carolina: Information Age Publishing.

Cole, R. A. (2000). *Issues in web-based pedagogy: A criticial primer.* Westport, CT: Greenwood Press.

Gordon, E. E. (2005). *Peer tutoring: A teacher's resource guide.* Lanham: R&L Education.

Hassan, F., & Selamat, N. F. (2002). *Why aren't students proficient in ESL: The Teachers' Perspective*. Retrieved 30 August, 2014, from MELTA: http://www.melta.org.my/ET/2002/wp10.htm

Kapp, K. M. (2012). *The gamification of learning and instruction: Game-based methods and strategies for training and education.* New York: John Wiley & Sons.

MacCarthy, M. (1998). *Spoken language and applied linguistics.* Cambridge: Cambridge University Press.

Mahavera, S. (2014). job seekers with a in spm english but can't speak a word of it. Kuala Lumpur: The Malaysian Insider.

Mahavera, S. (2014). *Job Seekers with A in SPM English but Can't Speak a Word of It.* Kuala Lumpur: The Malaysian Insider.

McNamara, E. (2014). Teacher positive feedback. In E. McNamara, *Positive pupil management and motivation* (p. 18). London: Routledge.

Mishra, K. C. (2005). *Correction of errors in English.* New Delhi: Sarup & Sons.

Nunan, D. (2001). *Second language teaching and learning.* Newbury: House Teacher Development.

Oxford English Dictionary. (2009). Oxford: Oxford University Press.

Parnell, A. (1989). Liaison interpreting as a language teaching technique. In L. Gran, & J. Dodds, *Theoretical and practical aspects of teaching conference interpretation* (pp. 253 –256). Udine: Campanotto.

Pascoe, M., Stackhouse, J., & Wells, B. (2006). Explaining intelligibility. In M. Pascoe, J. Stackhouse, & B. Wells, *Persisting speech difficulties in children* (p. 279). New York: John Wiley & Sons.

Rizvi. (2005). Improving fluency and self-expression. In Rizvi, *Effective tech communication* (p. 97). New York : Tata McGraw-Hill Education.

Robinson, M., Galaczi, E. D., Docherty, C., King, A., & Khalifa, H. (2014). Supporting National Education Reform: The Cambridge Malaysia Baseline Project, In *Research Notes: Issue 58* (pp. 40 – 43). Cambridge: Cambridge English Language Assessment.

Spradley, J. P. (1979). Universal Semantic Relationships. In J. P. Spradley, *The ethnographic interview* (p. 135). New York: Holt, Rinehart and Winston.

Ur, P. (1996). *A course in language teaching.* Cambridge: Cambridge University Press.

WadensjÖ, C., Dimitrova, B. E., & Nilsson, A.-L. (2007). *The Critical Link 4: Professionalisation of interpreting in the community: Selected papers from the 4th International Conference on Interpreting in Legal, Health and Social Service Settings, Stockholm, Sweden,* Amsterdam: John Benjamins Publishing Company.

Wells, J. C. (2006). *English intonation*. Cambridge: Cambridge University Press.