COMPARING FORM FOUR MALAY AND CHINESE STUDENTS' SPOKEN ENGLISH

Amerrudin Abd Manan
Sarimah Shamsudin
Universiti Teknologi Malaysia

ABSTRACT
This research was carried out to compare form four Malay and Chinese students' English language speaking ability. Data for the study was obtained from *The English of Malaysian School Students* (EMAS) corpus (Arshad et al., 2002). Among others, EMAS corpus comprises form four Malay and Chinese students' recorded oral description of a picture series. Based on their oral descriptions of the picture series, WordSmith Tools (Scott, 2008) was used to gauge both ethnic groups' language productivity. The software called RANGE (Nation, 2002) was also used to gauge their vocabulary range and sophistication. The results of the study show that the Chinese students performed better in terms of language productivity, vocabulary range and sophistication. This means that they uttered more words, used wider and more difficult vocabulary than the Malay students during the oral description. Comparing the achievements of the different ethnic groups in terms of spoken English can be useful for healthy competition. It helps groups that lag behind learn from those which are ahead. It can also stimulate research on why students from some ethnic groups perform better than others in spoken English for diagnostic and remediation purposes.

Keywords: spoken English, language productivity, vocabulary range, vocabulary sophistication

Introduction

Malaysian students are made up of many ethnic groups, and the three biggest are the Malays, Chinese and Indians. There have been studies comparing the three ethnic groups' performance in Mathematics and Science at the school level (Liew & Pong, 2005), graduates' CGPA scores at the Faculty of Business and Accountancy at University Malaya (Alfan & Md Nor, 2005) and the performance in Economics at the School of Accountancy and Business at Nanyang Technological University (NTU) in Singapore (Tay, 1994). Likewise studies

have also been conducted to see secondary school students' achievements in language learning (Hairuzila & Rohani, 2008; Rosna & Sharifah, 1994). Among Malaysians, there are opinions that Indian and Chinese students are better than Malay students in the English language, especially spoken English. In showing his concern on the achievement of Malay undergraduates' academic quality, Royal Professor Ungku Aziz, the former University of Malaya Vice-Chancellor once said, "…what was more disappointing was that their command of English was poorer compared to the other races" (The Star, September, 2000, p.16). These perceptions, however, would remain as mere opinions and speculations until they are backed up by hard evidence from research findings. Comparing the abilities of the different ethnic groups in learning English in our country can be useful. Firstly, it can promote healthy competition among students; secondly, the different ethnic groups can learn from each other and thirdly, the results from the comparison can stimulate teachers and researchers to investigate why some ethnic groups tend to perform better in learning English than others for diagnostic and remediation purposes.

The aim of this research is to compare the performance of Malay and Chinese students in spoken English at the secondary school level. Spoken English was chosen for the study instead of listening, reading and writing, because it is, arguably, the most important among the four language skills in an increasingly globalised world today. Its importance can perhaps be seen from the grouses that Malaysians have of the skill in question. One of the grouses is carried by a local newspaper (The Sun, July 2012, p.8) which reads, "Dissatisfaction with the level of spoken English in Malaysia has come to a boil, and every day a newspaper headline spotlights another aspect of the problem." Its importance can again be seen from the result of an online vote, carried out by the BBC, asking English teachers to indicate what they think is the most important aspect of Business English. The result seems to support our view on the growing importance of English today. From the total of 1151 votes, social English obtained 9.6% votes, writing business letters 14.2% votes, business vocabulary 15.7% votes, listening skills 25.5% votes and finally speaking skills obtained the highest votes, i.e. 34.9% (BBC Learning English for Teachers, n.d.).

There are but few related studies on this topic. Rahil, Habibah, Loh, Muhd, Nooreen and Abdullah (2006) compared the self-efficacy (the beliefs which individuals hold about their abilities and outcome of their efforts) of form four Malay, Chinese and Indian students in terms of learning English in general. Their study suggests that Indians have higher self-efficacy than Malay and Chinese students, and Malays have higher self-efficacy than Chinese students. Hairuzila and Rohani (2008) carried out a study to investigate a Malaysian private

university students' self-efficacy in spoken English. The results suggest that Indian students had the highest level of self-efficacy in spoken English followed by Chinese and then Malay students. In the above studies, Indian students seem to have the highest self-efficacy in spoken English while there is no conclusive finding on whether Chinese or Malay students have a higher self-efficacy.

The studies above are useful in providing us with insights on the performance of Malaysian secondary school students from the various ethnic groups in spoken English. Unfortunately, none of the research above looked at the actual performance of the students in their English language speech production. Rahil et al. (2006) and Hairuzila and Rohani (2008) obtained their findings from survey questionnaires that investigated their subjects' self-efficacy in learning English instead of the actual product or the oral linguistic output of the students. To better determine the speaking ability of the respective ethnic groups, it is felt that the actual linguistic output of students in their speaking ability should be obtained instead. To this end, researchers can make use of language corpora to investigate school students' oral data. Renouf (1987, p.1) defines a corpus as "a collection of texts, of the written or spoken word, which is stored and processed on computer for the purpose of linguistic research." Corpora (plural form of corpus) have been used widely by language instructors and linguists to analyse the structural patterns and typical features of real language in use.

In the Malaysian context, the use and analysis of English language corpora has been somewhat limited (Arshad, 2004). To our knowledge, the only spoken corpus collected from Malaysian secondary schools that is available is *The English of Malaysian School Students* (EMAS) corpus (Arshad, Fauziah, Jayakaran, Ghazali, Sharifah, Juridah & Edwin, 2002). We were therefore left with no other choice but to use the EMAS corpus for our investigation.

Methods

To compare the performance of Malay and Chinese secondary school students in their English speaking ability, the EMAS corpus (Arshad et al., 2002), compiled by researchers from Putra University of Malaysia was used. The EMAS corpus is a collection of written essays and speech production of primary five, form one and form four students. This corpus contains close to half a million words. The students were selected from three states, i.e. Penang, Melaka and Pahang, in an attempt by the compilers to represent the northern, the southern-central and the eastern parts of peninsular Malaysia. Specifically, this research will only investigate the speech production of form four students from the corpus. The form four students were chosen for the study because they spent the longest time in terms of basic primary and secondary education in the corpus. Compared to the primary five and form one students in the EMAS corpus, they would

therefore provide the best data in terms of spoken English. The different levels of the students' proficiency, ideally, should have been controlled at the beginning of the study. Comparing form four Malay and Chinese students of similar proficiency, for instance, those who scored A for PMR English would have added more weight to the results. In this study, however, the intention was only to base the study on a single constant variable, i.e. the roughly equal time they spent learning English at school.

The speech production data of the form four students is based on a picture series (see Appendix A) which shows an incident in which two girls were shouting to three boys on a fishing trip to help one of them who had fallen into a lake. Using the picture as a stimulus, the students were simply asked to narrate what they see, which might have taken them roughly between three to five minutes. As the stimulus was in the form of a series of pictures, the students did not require questions or prompts to get them to speak. The speech production of each student was recorded, transcribed and then saved in an individual digital file.

The students who took part in the study were selected from three states, i.e. Penang, Melaka and Pahang. Three schools were selected from each state. Because the recording and transcription of the students' speech was too time consuming, only 73 form four students were randomly chosen by the corpus compilers to take part in the oral narration. In this corpus, there are 33 Malay, 32 Chinese and 8 Indian students' oral data, but since this study aimed to make a comparison between Malay and Chinese students only, data from the Indian students were removed. Since the oral data files in the corpus were somewhat limited, all the remaining 33 Malay and 32 Chinese students' oral data were collected and used for the study based on convenient sampling method. To balance the number, one data file from the Malay students was thus randomly removed from the corpus.

To compare the oral performance of the Malay and Chinese students, totalling 64 in this study, we carried out an investigation of each group in two major linguistic areas: language productivity and vocabulary. Such a method was used by Arshad (2004) to examine the English language development of students from primary to secondary school. An investigation on our subjects' language productivity involved a study on the number of words and sentences the Malay and Chinese students produced while an investigation on their vocabulary involved a study on the range and sophistication of the vocabulary of each group during the oral narration. It is conceded, however, that the measurement used to gauge the oral performance of both groups of students was somewhat limited, because we only had access to the transcribed version of the students' oral

narration. It was not anywhere close to the measurement of oral skills as that used for *SPM (Sijil Pelajaran Malaysia*) or MUET (Malaysian University English Test) oral examinations. Evaluating the students' oral performance from their actual speech production and that which had been transcribed may perhaps give rise to different findings. To gauge the language productivity of each group, the WordSmith Tools (Scott, 2008) software was used. All oral data files of the Malay and Chinese students numbering 32 each were run using WordSmith Version 5.

Results

Language Productivity

Table 1 illustrates the total number of sentences and words for each group and the total number of sentences and words for each Malay and Chinese student.

Table 1: The students' language productivity

Ethnic group	Total number of sentences	Average sentences per student	Total number of words	Average words per student
Malay	250	8	3098	97
Chinese	399	12	5216	163

It can be seen from Table 1 that the Chinese students were more productive in terms of the number of words and sentences produced during the oral narration.

Vocabulary Range

The range or diversity of vocabulary can be measured, according to Schmitt (2002), by calculating the type to token ratio in a corpus. The term *token* according to Hunston (2002) refers to the occurrence of words in a text regardless of whether the words are repeated or otherwise. The term *type*, according to Hunston (2002), however, refers to words which are different or not repeated in a text. Type to token ratio is also known as lexical density (Richards et al., 1985). According to Richards et al. (1985), it is normally expressed in the form of percentage and is calculated by the formula below:

$$\text{Type to token ratio (Lexical density)} = \frac{\text{Number of separate (different) words}}{\text{Total number of words in the text}} \times 100$$

As uncommon words tend to be used less frequently than common words, a smaller type to token ratio obtained from a user would suggest that the user's

language is less diverse as s/he uses only a limited range of vocabulary. By contrast, a larger type to token ratio obtained from a user would suggest that the user's language is more diverse as s/he uses a wider range of vocabulary. Type to token ratios of the two groups of students were obtained by running the data files with the software called RANGE (Nation, 2002). Table 2 illustrates the type to token ratios of the two groups of students.

Table 2: The students' type to token ratio

Ethnic Group	Types	Token	Type/Token Ratio
Malay	475	3131	15.17
Chinese	815	5276	15.44

It can be seen from Table 2 that the Chinese students used a larger number of word types than the Malay students, an indication that they had a more diverse or wider vocabulary range. Nevertheless, the closeness of the ratios between the two groups suggests that the Chinese students' vocabulary was not as highly diverse as one would expect it to be. Although they produced a large number of words during the oral narration, the ratios suggest that they only produced a relatively slightly larger amount of diverse vocabulary than the Malay students.

Vocabulary Sophistication

Vocabulary sophistication of the two groups of students can also be gauged using RANGE (Nation, 2002). Besides calculating word type to token ratio, RANGE can also categorise the word types that the Malay and Chinese students used in their speech production. RANGE categorises three lists of word types based on their frequencies. The first is categorised under base list one, the first most frequent 1000 words group. The second is categorised under base list two, the second most frequent 1000 words group. The third is categorised under base list three, i.e. words which do not appear in the first two groups. According to Nation (2002), the source of the first and second category of words is the General Service List of English Words compiled by West (1953), and the source of the third category of words is the Academic Word List compiled by Coxhead (2000). In terms of vocabulary sophistication, therefore, the most sophisticated would be those in the third category, as they are the least frequent. In short, whereas vocabulary range or diversity describes the width or breadth of a person's vocabulary, vocabulary sophistication describes a person's mastery of difficult vocabulary. Tables 3 and 4 show the distribution of word types used by the two groups of students based on the three word lists.

Table 3: Frequency and percentage of word types and word families of Malay students according to categories

Base list	Types		Word families
	Frequency	Percentage	
One	332	69.89	246
Two	69	14.53	55
Three	10	2.11	10
not in the lists	64	13.47	-
Total	475	100	-

Table 4: Frequency and percentage of word types and word families of Chinese students according to categories

Base list	Types		Word families
	Frequency	Frequency	
One	509	62.45	360
Two	127	15.58	98
Three	26	3.19	24
not in the lists	153	18.77	-
Total	815	100	-

It can be seen from Tables 3 and 4 that the Malay students have a higher percentage of word types in base list one compared to the Chinese students. This shows that the Malay students tended to rely more on the most frequent 1000 words group than the Chinese students during the oral narration. In contrast, it can also be seen that the Chinese students have a bigger percentage of word types in base list three group compared to the Malay students. This indicates that the Chinese students have a larger amount of vocabulary in the least frequent or the more difficult Academic Word List than the Malay students. Tables 3 and 4 thus show that the Chinese students' vocabulary is more sophisticated than the Malay students. As can be seen in Appendix B, the amount of academic vocabulary or base list words produced by the Chinese students is more than twice the amount produced by the Malay students in the EMAS corpus.

Discussion and Conclusion

This study makes a comparison between Malay and Chinese form four secondary school students in their English language speaking ability. The discussion will, however, also be compared to college and university students as well, as it is felt

that the English language learning variables of Malay and Chinese students in secondary and tertiary level institutions, arguably, would not greatly differ.

Compared to the study by Hairuzila and Rohani (2008), for example, our study makes an attempt to enhance the validity in the measurement of the students' spoken English. While they used a survey questionnaire to compare the students' oral ability from the self-efficacy aspect, the present study made use of actual speech production of the students in an oral narration obtained from the EMAS corpus. When examining the results of the study in terms of language productivity, vocabulary range and sophistication, it was found that the Chinese students fared better than the Malay students in all the three aspects mentioned. In other words, the Chinese students, during the oral narration, not only uttered more words, but they also possessed a wider vocabulary and had mastery of more difficult vocabulary than the Malay students. This research therefore suggests that for spoken English, the Chinese students were better than the Malay students in terms of language productivity, vocabulary range and sophistication.

Though it is acknowledged that there are other variables which impact the learning of English, when considering these findings, however, they are rather perplexing because the two ethnic groups spent about the same amount of time learning English in school. By right, their performance should be roughly the same. The reason for this disparity is not known, because we only attempted to investigate the performance of the two groups of students' spoken English in terms of language productivity, vocabulary range and sophistication. Investigation on the cause of this disparity is beyond the scope of this study. Nevertheless, a number of reasons that may contribute to this disparity could still be offered, albeit tentatively, based on our long experience as English language teachers, the demographic information that is available in the EMAS corpus and the findings from related research. First, a look at the students' demography might to a certain extent, offer one explanation for this disparity. While a number of the Chinese students in the corpus reported that they used English as the language they speak at home, none of the Malay students reported so. This may suggest that the Malay students were less motivated to speak in English and also lacked practice in speaking the language. In contrast, the Chinese students' use of English at home suggests that they were more motivated in learning to speak English and had more practice in speaking the language.

On the issue of speaking in English especially outside the classroom, the dilemma faced by the Malay students is quite well-known. Malay students might be more distant from the English language than their Chinese counterparts, as

many still feel that when they speak in English, fellow Malays may see them as being too westernised or that they are trying to show off (Thang, 2004). This perception could also be due to the resentment that some Malays have towards British colonisation in the past (Harmi, 2007). A study carried out by Mardziah and Wong (2006) could perhaps lend weight to the suggestions above. Mardziah and Wong carried out a survey to study the attitude of Malaysian students towards English language learning and how the language impacts on the students' ethnic and national identity. The subjects of their study were 331 university and secondary school students from different ethnic groups residing in peninsular Malaysia. The results of the study suggest that Indian students have the most positive attitude towards the English language and learning the language. Chinese students and the ethnic group, labeled 'Others' (comprising natives from Sabah and Sarawak) came second and third respectively. The Malay group, however, was the group that most strongly perceived English and learning the English language as a threat to their ethnic and national identity. The study also found that Indian students were the most comfortable while the Malays the least comfortable when using English in communication. In terms of fluency in speaking English, Indian students were found to be most proud of this achievement compared to the other three ethnic groups. Also, while Indian, Chinese and the other ethnic groups would seek approval and recognition in their English language achievement from friends and family members, the study found that such sentiment was not present among the Malay students.

The findings of another study suggest a positive relationship between higher proficiency level and positive attitude and motivation to learn English (Thang, Ting & Nurjanah, 2011), indirectly implying that those with higher proficiency tend to be more successful than those with lower proficiency. Our study suggests that because Chinese students perform better in spoken English than Malay students, they may have more potential for future success than the Malay students. For example, in local colleges and universities involving academic activities today, there is an increasing demand to speak in English, and "speaking fluent English is an important criterion for graduates to succeed as professionals…" (Dhillon, 2011, p.12). In English language classes which are often compulsory for students in these institutions, the Chinese students' English speaking advantage would enable them to participate actively in those classes.

Besides English classes, it is not uncommon to find university subject lecturers delivering their lectures in English today. With the big increase in the number of foreign students entering our tertiary institutions today, the use of English in all the four skills looks set to become the norm rather than the exception. If students are less proficient in spoken English, it may limit their participation and increase their frustration in those classes too. Though it may be too early to predict,

Chinese students may also have better chances in terms of securing jobs after they have completed their education because of their oral English advantage. As the saying goes, success breeds success.

The results of yet another research by Rahil, Noran and Habibah (1994) would help English teachers further understand why students from some ethnic groups perform better than others in learning English. Rahil et al. (1994) used an instrument called STAI (State-Trait-Anxiety Inventory) to gauge the anxiety level of 1215 first year Putra University of Malaysia students from different ethnic groups taking compulsory English courses at the university. Their research results showed that during English classes, the 1215 students' anxiety level was highest for speaking, followed by writing, comprehension and reading respectively. In terms of ethnicity, the research indicated that anxiety level was highest among Malay, Chinese and Indian students respectively. By place of origin, the research showed that the anxiety level was highest among students from rural areas, followed by those from small towns and big towns respectively. Male students' anxiety level was also found to be higher than that of female students. Since there may be a tendency for some Malaysian teachers to teach English without considering the traits, personality and hardships that their students go through, research such as this may shed light on why some ethnic groups fare worse than others in mastering English language skills.

It is thus clear that the Malay students should make more effort to improve their English speaking ability to be on par, if not better than their Chinese and Indian counterparts. They should keenly observe and emulate the learning strategies of Chinese and Indian students to improve their speaking ability. There is no harm in adopting the strategy of the Chinese and Indian students, i.e. practising to speak the language outside the classroom and at home. If they wish to polish up their speaking ability, negative perceptions and attitudes such as shyness and worry about making mistakes when speaking need to be eliminated. For their part, Malaysian English teachers at secondary and tertiary education may have to come to terms with the fact that most Malay students do not have the chance to speak English outside their schools especially those living in rural areas. Providing real-life situations for them to practise speaking in English in the classroom using meaningful communicative approach activities is therefore crucial to Malay students.

There is a need for further research on this topic to help some groups in our country who are not doing as well as others in mastering English language speaking skills that are growing in importance today. Weakness in English among any of the ethnic groups in the country could lead to dispossession in

some spheres of life, such as learning, business, travel and employment opportunities. Mastering English language speaking skills can indeed empower our people to become more functional and productive in many spheres of life in the local and globalised arenas.

References

Alfan, E., & Md Nor Othman (2005). Undergraduate students' performance: the case of University of Malaya. *Quality Assurance in Education*, *3*(4), 329-343.

Arshad, A.S., Fauziah, H., Jayakaran, M., Ghazali, K., Sharifah, Z. S. A. R., Juridah, M. R., & Vethamani, E.M. (2002). *The English of Malaysian School Students (EMAS) Corpus*. Serdang: Universiti Putra Malaysia.

Arshad, A. S. (2004). Beyond concordance lines: using concordances to investigating language development. *Internet Journal of e-Language Learning & Teaching, 1*(1), 43-51.

BBC Learning English for teachers – Your ideas (n.d.). Retrieved from www.bbc.co.uk/worldservice/learningenglish/teachingenglish/yourideas/teaching_busi-ness_english.shtml.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*, 213-238.

Dhillon, P.K. (2011). Using learning styles to enhance ESL acquisition. *The English Teacher*, *40*, 11-26.

Hairuzila, I., & Rohani, S. (2008). Perceived self-efficacy of Malaysian ESL Engineering and Technology students on their speaking ability and its pedagogical implications. *The English Teacher*, *37*, 61-75.

Harmi, I.B., Tretiakov, A., & Kinshuk (2007). Teaching business English to adult Malay learners: the potential of agent technology. *Proceedings of Ascilite, Singapore*, December.

Liew, H.P., & Pong, S.L. (2005). *Mathematics and Science achievements in Malaysia*. Retrieved from http://paa2005.princeton.edu/download.aspx?submissionId=50599

Hunston, S. (2002). *Corpora in Applied Linguistics.* Cambridge: OUP.

Mardziah, H.A. & Wong, B.E. (2006). Listening to the ethnic voice in ESL learning. *The English Teacher, 35*, 15-26.

Nation, P. (2002). *Range and frequency: programs for Windows based PCs.* Retrieved from http://www.victoria.ac.nz/lals/about/staff/paul-nation

Rahil, H.M., Noran, F.Y., & Habibah, E. (1994). Anxiety towards learning English among Universiti Pertanian Malaysia Students. *Akademika, 45*, 57-67.

Rahil, H.M., Habibah, E., Loh, S.C., Muhd, F.M., Nooreen, N. & Abdullah, M.C. (2006). The relationship between students' self-efficacy and their English language achievement. *Jurnal Pendidik dan Pendidikan, 21*, 61-71.

Renouf, A. (1987). Corpus development. In J. M. Sinclair (Ed.), *Looking up: an account of the COBUILD Project in lexical computing* (pp. 1-40). London: Collins COBUILD.

Richards, J., Platt, J., & Weber, H. (1985). *Longman dictionary of Applied Linguistics.* Essex: Longman.

Rosna, A. H., & Sharifah, A. S. S. (1994). Examining learners' language learning strategies. *RELC Journal, 25*(2), 1-20.

Schmitt, N. (2002). Using corpora to teach and assess vocabulary. In M. Tan (Ed.), *Corpus studies in language education* (pp. 31-44). Thailand: IELE Press.

Scott, M. (2008). *WordSmith tools Version 5*. Liverpool: Lexical Analysis Software.

Tay, R.S. (1994). Students' performance in Economics: Does the norm hold across cultural and institutional settings? *The Journal of Economic Education, 25*(4), 291-301.

Thang, S.M. (2004). Learning English in multicultural Malaysia: Are learners motivated? *Journal of Language and Learning, 2*(2), 142-153.

Thang, S.M., Ting, S.L. & Nurjanah, M. J. (2011). Attitudes and motivation of secondary students towards learning English as a second language: a case

study. *The Southeast Asian Journal of English Language Studies, 17*(1), 40-54.

The Star (2000). Dr. M: Where have the Bumi male students gone to? 4 September: p. 16.

The Sun (2012). Mastering the language of business. 2 July: p. 8.

West, M. (1953). *A general service list of English words, with semantic frequencies and a supplementary word-list for the writing of popular Science and Technology*. London: Longmans Green.

APPENDIX A
Picture Series

Adopted from: Arshad et al. (2002). *The English of Malaysian School Students (EMAS) Corpus*.

APPENDIX B
List of Academic Word Types

No	Chinese	Malay
1	accompanied	approached
2	accompanying	classical
3	aid	equipment
4	apparently	ignore
5	appreciation	incident
6	area	labeled
7	assembly	location
8	declined	medical
9	depressed	stable
10	equipment	topic
11	eroded	
12	finally	
13	ignored	
14	incident	
15	injured	
16	partner	
17	principal	
18	project	
19	ranging	
20	reacted	
21	seeking	
22	selected	
23	series	
24	stabilised	
25	unstable	