

Chiu-yu Tseng

Institute of Linguistics, Academia Sinica, Taiwan

cytling@sinica.edu.tw

ABSTRACT

This paper reports an on-going multinational collaboration “Asian English Speech cOrpus Project” (AESOP). The aim of the consortium is to build up an open resource English speech corpus representing the varieties of English spoken in Asia. It was agreed in 2008 that each participating member will use the same recording platform to collect a common set of core data so that a set of core properties common to all varieties of Asian English, as well as to features that are particular to individual varieties can be developed. Areas of overlap between L1 and L2 English and differences among varieties of Asian English could be determined. Initial efforts in 2009 were devoted to designing spoken-language tasks that aims to elicit the phonetic aspect of a large range of English segmental and suprasegmental characteristics. The 2011 efforts, reported here by the Taiwan members, were devoted to designing spoken language tasks that aims to elicit (1) a comprehensive coverage of phonotactic variations, (2) broad vs. narrow focus and (3) dialogue and discourse materials using discourse completion tasks (DCT). We believe the data collected can be used to develop ASR tools, investigate various issues of speech prosody, and assess English proficiency towards CALL applications.

Index Terms—Asian L2 English, phonotactics, broad and narrow focus, discourse completion tasks, speech prosody language proficiency assessments

1. INTRODUCTION

The TaiWaN Asian English Speech cOrpus Project (TWNAESOP) is part of the ongoing multinational collaboration “Asian English Speech cOrpus Project” (AESOP) whose aim is to build up a consortium of English speech corpus. Each research team will use a common recording setup and share an experimental task set, and will develop a common, open-ended annotation system. AESOP-collected corpora which represent many varieties of English spoken in Asia will be an open resource, available to the research community at large. In 2009, the TWNAESOP team presented data design to be used as core content for all AESOP collaborators so that the consortium could construct databases with a small portion of overlapping content for commonality investigations as well as additional

data to suit individual needs [1]. The 2009 AESOP core design focused on eliciting production of a large range of segmental, suprasegmental characteristics from 6 read speech tasks and 2 spontaneous speech tasks, namely, (1) Target Words in Carrier Sentences (2) Target Words at Phrase Boundaries (3) Target Words in Contrastive Stress Positions (4) Stressed and Unstressed Function Words (5) Prosodic Disambiguation (6) Reading a Passage “The North Wind and the Sun” (1) Computer-Prompted Dialogue (2) Picture Description Task [1, 2, 3]. Since the Taiwan AESOP research team (TWNAESOP) which we represent were among the first AESOP members to obtain research funds (TaiWaN Asian English Speech cOrpus Project, Chiang Chih-Kuo Foundation for International Scholarly Exchange (CCKF) project number DB002-D-08, 2009.7.1-2012.6.31), we were able to start data collection from the fall of 2009. Speech data of the core content from 500 speakers were successfully collected in 2010, including 12 American L1 speakers and 488 Taiwan Mandarin speakers [4].

However, while the controls used for the core design included the phonetics aspects of English segmental and supra-segmental in one hour of recording time per speaker, the collected data did not include sufficient phonotactic variations or a wide range of prosodic and pragmatic aspects of L2 English as well.

In order to collect more speech data to remedy the above mentioned problems, we have designed a different set of materials. A review of some of the major corpora including TIMIT, IViE, UME-ERJ and CASSAESOP revealed the following picture. The 1993 TIMIT corpus of read speech was designed to provide speech data for acoustic-phonetic studies and for the development and evaluation of automatic speech recognition systems [5]. Its content includes read speech of (1) 460 phonetically-compact sentences (2) 2 calibration sentences and (3) 1890 randomly selected sentences [6]. The corpus has been widely used by the speech science community since its appearance and proved to be quite suitable for general acoustic-phonetic research. The 1998 IViE corpus by the Phonetics Lab, University of Oxford is designed to investigate intonation variation of British English. Its contents includes read speech of (1) 22 phonetically-controlled sentences with different grammatical structures sentences and (2) a passage of “The Cinderella Fairy Tale”; and spontaneous speech of (3) a retold version of the Cinderella fairy tale (4) a map task and (5) discussion on a

given topic [7]. The 2000 UME-ERJ (Utilization of Multimedia to Promote Higher Education Reform-English Read by Japanese) corpus of read speech by Tokyo University is designed to support CALL (Computer Assisted Language Learning) related research through collecting segmental and suprasegmental speech materials read by Japanese students. The content of the UME-ERJ corpus design includes 8 sentence sets and 5 words sets. The sentence sets consist of the TIMIT phonetically compact sentences, single intonational phrases, sentences with intonation markings and sentences with stress markings. The word sets consist of list of single words, short phrases and compound words [8, 9]. The 2010 CASSAESOP corpus by The Chinese Academy of Social Sciences is by far the largest in size and most comprehensive in content materials. It consists of all of the phonetically compact sentences from TIMIT, the Cinderella Fairy Tale from IViE, the text pieces from UME-ERJ, the AESOP core sentences [1], and two additional sections of their own design, namely the CASS-E (English) and CASS-C (Chinese). The CASS-E set was specially designed to cover a large variety of intonation variations.

Our goal is to collect sufficient phonotactic variations as well as more prosody and discourse data, while facilitating the AESOP goal of open resource and data sharing at the same time. We discussed with the CASSAESOP team as well as the Japan AESOP team at Waseda University and reached a consensus to collect some common data of mutual interests as well. As a result, the TWNAESOP2 project reported below is set to collect an entirely different set of speech data with the following goals: (1) a large coverage of phonotactic variations through read speech of isolated words (2) the passage of "The Cinderella Fairy Tale" that is considerably longer than "The North Wind and the Sun" used before (3) the CASSAESOP designed broad/narrow focus sentences and (4) the Waseda developed DCT (Discourse Completion Tasks) discourse data through elicited dialogues.

In the following sections, we will report how we designed a set of phonotactic variations (Sec. 2.1.), what we adopted from CASSAESOP (Sec. 2.2; 2.3), and how we tailored the Waseda DCT to suit Taiwan speakers (Sec. 2.4). The speech data we collect will be termed TWNAESOP2

2. CORPUS DESIGN

2.1. Phonotactic Variations from Phonetically-rich Isolated Words

To generate a large coverage of phonotactic variations, the design rationale of the PhoneBook which aims to design a set of word lists that cover most if not all significant coarticulatory variants in a large variety of phonetic contexts was adopted [10]. The PhoneBook developed a two-stage process to generate word lists from the CMU electronic dictionary. At the first stage, a list of candidate words is

generated through a manual filtering process that deletes unsuitable words to avoid reading difficulties that speakers may experience. Foreign words, difficult and obscure words, words with multiple acceptable pronunciations, potentially embarrassing words, likely misread word, homonym pairs and acronyms are deleted through a filtering process. The phonemic contexts, such as tri-phones and syllable-based templates, can therefore be enumerated from candidate word list. At the second stage, the final word list is generated by choosing a subset of candidate word lists that captures each enumerated context at least once. The frequency of phonemic contexts is used as reference of scoring estimation. Following the PhoneBook design rationale but with automatic instead of manual filtering, the following open resources were used for the present project, namely, (1) the syllabified CMU electronic dictionary [11] and (2) part of speech (POS) from a word frequency list [12]. The filtering process is reported in Sec. 2.1.2.

2.1.1. Source Materials

2.1.1.1. Syllabified CMU Pronouncing Dictionary (version: 0.6)

The syllabified CMU Pronouncing Dictionary (version: 0.6, 2009) is used for the present study. Syllable boundaries (s) were generated automatically of word entries in the CMU Pronouncing Dictionary using a structured SVM approach. It is reported that Over 98% of the words were correctly syllabified [11].

2.1.1.2. The Word Frequency List

We also used an open resource of word frequency to help filter frequency information. The on-line resource [12] contains approximately 500,000 words with information of word frequency, word form, part of speech (POS), and the number of texts in which any given word occurs. In addition, each word appears at least four times in the Corpus of Contemporary American English (COCA, 410 million words).

2.1.2. Methods

Both the syllabified CMU dictionary and word frequency list were used to filter unsuitable words specified by the PhoneBook. In addition, we also removed words which already appeared in TWNAESOP1 and other texts adopted by TWNAESOP2. A new candidate word list is generated as a result. The same algorithm of word list generator of the PhoneBook was then applied to the frequency of phonemic contexts.

2.1.2.1. Word Filtering

The following procedures show how word filtering is derived:

1. Adopting the CMU Pronouncing Dictionary as our reference to remove misspelled words.
2. Deleting homonym pairs and acronyms by phone

sequence of the CMU dictionary.

3. Deleting words including following noun of title preceding, noun of title locative noun and proper noun in the POS tagging.
4. Deleting existing words which appeared in TWNAESOP1 and TWNAESOP2.

2.1.2.2. Word List Generation

Phonemic context was enumerated from the candidate word list. According to the PhoneBook, derived tri-phone candidates and syllable-based templates yielded similar coverage of contexts. Therefore, we only used the syllable-based template as phonemic context. The candidate word list and enumerated phonemic context were used as input of the word-list generator. The procedure is shown as follows.

1. Counting frequency of enumerated phonemic context in the candidate word list
2. Generating candidate words whose scores are reciprocal of enumerated phonemic context of word frequency. Words of higher scores represent those with unique phonemic contexts.
3. Selecting words by the highest score as output word
4. Rescoring the remaining unselected candidate words
5. Repeating steps 1 to 4 until convergence frequency of enumerated phonemic context is reached.

2.1.3. Results

A 32,950-word candidate list was derived after word filtering (Sec. 2.1.2.1.), which was reduced to a list of 6,122 words after word list generation (Sec. 2.1.2.2).

2.2. Broad and Narrow Focus

A total of 415 sentences were used by the CASSAESOP to collect a large variety of intonation variation, including the 22 IViE phonetically-controlled sentences with different grammatical structures sentences. We selected a total of 176 sentences (1784 words) that were designed specifically to elicit broad and narrow focus which can be used to help our other studies on prosodic highlights such as prominence and focus in fluent continuous speech. A pilot recording session of the 176 sentences by one Taiwan speaker showed that approximately 90 minutes of recording. Examples can be found in Appendix A.

2.3. A Longer Passage “The Cinderella Fairy Tale”

The IPA (International Phonetics Association) has provided slightly different versions of “The North Wind and the Sun” story to cover the entire set of phonemes in English. The text we adopted for TWNAESOP [1] yields a total of 113 words (144 syllables) in 3 paragraphs of 8 independent clauses and 5 dependent clauses. The recording time to read the passage ranged from 40 to 50 seconds [13]. While it is an easy passage for speakers and takes little time to record,

it does not provide sufficient intonation variations or discourse prosody. On the other hand, the passage of “The Cinderella Fairy Tale” chosen by IViE corpus is considerably longer and consists of more varied intonations and discourse prosody in read narrative. The text yields a total of 759 words (1,000 syllables) in 14 paragraphs of 93 independent clauses and 49 dependent clauses. Test recording time for one L2 speaker was approximately 5 minutes and 30 seconds. The passage is included in the content of CASSAESOP, so we, too, decided to use the text used from CASSAESOP to in order to collect comparable narrative data. (See Appendix B)

2.4. Discourse Completion Tasks (DCT)

The DCT is a well-established and widely adopted tool to investigate speech acts such as request and refusal performances in second language learning [14]. The aim of the design is to obtain more interactive discourse speech data by using the Discourse Completion Task as an instrument. The Waseda University team has developed interactive conversation tasks through specified context that aims to elicit scenario appropriate questions and responses covering a wide range of personal and object domains [15]. A total of 86 2-speaker A-B dialogues 3 3-speaker A-B-C dialogues that are socially and culturally Japan appropriate was constructed. Five levels L2 English proficiency, from A2-basic to C2-advanced, was also built into the design for proficiency evaluation [16]. Each of the five levels contains 19 units and anywhere from 15~17 dialogues. We adopted the 86 2-speaker A-B dialogues (6,208 words) and tailored the social and cultural setting to be Taiwan appropriate.

Table 1 shows the distribution of the adopted 86 DCTs by L2 proficiency levels A2-basic to C2-advanced and size in the number of words by discourse, turn and sentences.

Table 1 DCT by L2 proficiency levels and number of words by discourse unit, speaker turn sentence.

| Proficiency Level # of words | A2 | B1 | B2 | C1 | C2 | Total |
|---------------------------------|---------|---------|---------|---------|---------------|-----------------|
| (A-B)/ Discourse | 12~85 | 38~96 | 42~114 | 59~108 | 64~197 | 12~197 |
| Turn-A/ sentence | 2~27 | 1~28 | 2~24 | 3~39 | 2~61 | 1~61 |
| Turn-B/ sentence | 2~37 | 2~35 | 1~45 | 6~41 | 1~38 | 1~45 |
| Total (A/B) | 447/411 | 519/539 | 662/601 | 607/671 | 1,023/ 728 | 3,258/ 2,950 |

A discourse is defined as a unit of dialogue which ranges from 12 to 197 words in size. The size of each turn/speaker ranges between 1 to 61 words for turn-A and 1 to 45 words for turn-B. The total number of words is 3,210 for turn-A and 2,906 words for turn-B. Examples can be found in Appendix C.

3. DATA CONTENT AND RECORDING PLATFORM

3.1. Data Content

The content of TWNAESOP2 is summarized in Table 2. The data set consists of 6,122 isolated words that are phonetically-rich, 176 short sentences (1,784 words in total) of both broad and narrow focus, 1 passage of “The Cinderella Fairy Tale” and 86 DCTs.

Table 2 Data content of TWNAESOP2

| Set | Token | # of L1/L2 speaker | Size/speaker (L1 /L2) | Total (L1/L2) |
|--|-------|--------------------|-------------------------------------|---------------|
| Phonetically-rich isolated words | 6,122 | 10/30 | 620/250 (20/50 overlapped words) | 6,200/7,500 |
| Broad focus and narrow focus sentences | 176 | | 176/88 | 1,760/2,640 |
| The Cinderella Fairy Tale | 1 | | 1/1 | 10/30 |
| Discourse Completion Tasks (DCT) | 86 | | 86/86 | 860/2580 |

Since the data size is many folds larger than the previous AESOP core content [1], we need to take recording time and fatigue factor into consideration. Our goal is record the speech data of 10 L1 speakers (5 males and 5 females) who will provide a larger coverage of the data set; and 30 L2 Taiwan speakers (15 males and 15 females) who will provide a smaller coverage using the TIMIT rationale to still yield a somewhat comprehensive coverage. In other words, each L1 speaker will be asked to provide one reading of 620 isolated words randomly selected from the list of 6,122 phonetically-rich words (overlapping 20 words per speaker), the 176 short broad-and-narrow-focus sentences, the passage of The Cinderella Fairy Tale, and the 86 DCTs. In turn, each L2 speaker will be asked to provide one reading of 250 isolated words randomly selected from the list of 6,122 phonetically-rich words (overlapping 50 words per speaker), 88 randomly selected broad-and-narrow-focus sentences from the original 176 sentences, the same passage of The Cinderella Fairy Tale and the 86 DCTs. When completed, we expect to collect L1 speech data of a total of 6,200 isolated words, 1,760 broad-and-narrow-focus sentences, 10 passages of “The Cinderella Fairy Tale” and 860 DCT dialogues. The L2 speech data will consist of 7,500 isolated words, 2640 sentences, 40 passages and 3,440 DCT dialogues.

3.2. Recording Platform

Following the AESOP protocol that each research team will use a common recording setup developed by the by the

Chinese University of Hong Kong, we will continue to use the TWNAESOP recording platform which was developed on the basis of the CUHK-SIAT recording tool and used for TWNAESOP data collection. Minor modifications are made in order to fit the requirements of TWNAESOP2.

3.2. Data Collection and Annotation

AESOP has agreed to design recording tasks to be carried out in naturalistic settings instead of soundproof chambers so that recording can easily be performed by research assistants in relatively quiet environments. The majority of the L1 speakers for TWNAESOP is still available and will be solicited. The L2 speaker population will consist primarily of undergraduate and graduate students. AESOP members also plan to develop a common protocol for file formats and annotation, and hopefully to share some preprocessed data such as forced alignment of segmental identities processed by the HTK Tool.

5. CONCLUSION

We have reported a second set of content design that aims to elicit a comprehensive coverage of phonotactic variations, expressing broad vs. narrow focus and dialogue acts through discourse completion tasks (DCT). Speech data from both L1 and L2 English speakers will be collected. The research is continuation of an on-going multinational collaboration that aims to build up an Asian English speech corpus representing the varieties of English spoken. During the initial stage (2009-2010), we have successfully collected speech data of the AESOP core content from 500 speakers from Taiwan [1, 4]. The data collected from spoken-language tasks can be used to investigate large range of English segmental and suprasegmental characteristics [4]. The content design reported in this paper, in turn, centers on phonotactic coverage, focus and prominence, narrative of considerable length and dialogue/discourse completion. We believe the data collected from the well selected isolated words could be effectively used towards speech technology development while the data collected from the broad/narrow focus design and reading out loud a long story could be used towards more understanding of narrative and discourse associative patterns, especially expressive prosody that highlights crucial information as well delivers speaker intension and emotion. Last but not least, the data collected from the DCTs will provide a wide range of social scenarios and cultural understanding that is as significant in L2 learning as learning the sound system, the words and grammatical structures. We believe the data collected can be used to facilitate a wide array of phonetic, prosodic investigations as well as testing materials for proficiency assessment, language education and applications in computer-aided language learning (CALL).

Acknowledgements

We would like to thank Professor Yoshinori Sagisaka of Waseda University who not only initiated the AESOP consortium in 2008 but also continued to lead the project till present. We would also like to thank Professor Helen Meng of the Chinese University of Hong Kong for developing the CUHK-SIAT recording tool and offering it as the common recording platform for AESOP. Professor Aijun Li and Dr. Yuan Jia of CASS (The Chinese Academy of Social Sciences) graciously provided us with their CASSAESOP data content and some speech data which helped us a great deal in our design to elicit broad/narrow focus. Finally, our special thanks go to Professors Michiko Nakano, Mariko Kondo and Yusuke Kondo for providing the DCT content and their invaluable suggestions of the DCT through endless e-mail communications and sessions of face to face discussion.

6. REFERENCES

- [1] T. Visceglia, C. Tseng, M. Kondo, H. Meng, and Y. Sagisaka, "Phonetic Aspects of Content Design in AESOP (Asian English Speech cOrpus Project)," *Oriental COCOSDA 2009*, Beijing, China, 6 pages, 2009.
- [2] T. Visceglia and C. Tseng, "Spontaneous Speech Database Design for Suprasegmental Characteristics in Asian L2 English," *NCMMSC 2009*, Urumqi, China, 6 pages, 2009.
- [3] H. Meng, C. Tseng, M. Kondo, A. Harrison and T. Visceglia, "Studying L2 Suprasegmental Features in Asian Englishes: A Position Paper," *Interspeech 2009*, Brighton, England, 4 pages, 2009.
- [4] C. Tseng, "Report on TAIWAN-AESOP," Presented at the *AESOP 2010 Workshop*, Katmandu, Nepal, 2010.
- [5] <http://www ldc.upenn.edu/Catalog/CatalogEntry.jsp?catalogId=LDC93S1>
- [6] V. Zue, S. Seneff, and J. Glass, "Speech Database Development at MIT: TIMIT and Beyond," *Speech Communication 1990*, North Holland, vol. 9 no.4, pp. 351-356, 1990.
- [7] E. Grabe and B. Post, "Intonation Variation in the British Isles," *Speech Prosody 2002*, Aix-en-Provence, France, 4 pages, 2002.
- [8] N. Minematsu, Y. Tomiyama, K. Yoshimoto, K. Shimizu, S. Nakagawa, M. Dantsuji, and S. Makino. "English Speech Database Read by Japanese Learners for CALL System Development," *Proc. LREC 2002*, Las Palmas, Spain, pp. 896-903, 2002.
- [9] N. Minematsu, Y. Tomiyama, K. Yoshimoto, K. Shimizu, S. Nakagawa, M. Dantsuji and S. Makino, "Development of English Speech Database Spoken by Japanese Learners," *Oriental COCOSDA 2001*, Taejon, Korea, 6 pages, 2001.
- [10] J.F. Pitrelli, C. Fong, S.H. Wong, J.R. Spitz, and H. C. Leung, "Phonebook: a Phonetically Rich Isolated Word Telephone Speech Database," *Proc. ICASSP 1995*, Detroit, U.S.A, 4 pages, 1995.
- [11] S. Bartlett, G. Kondrak and C. Cherry, "On the Syllabification of Phonemes," *Proc. HLT-NAACL 2009*, Colorado, U.S.A, pp. 308-316, 2009.
- [12] <http://www.wordfrequency.info/>
- [13] C. Tseng, Z. Su, C. Huang, and T. Visceglia, "An Initial Investigation of L1 and L2 Discourse Speech Planning in English." *ISCSLP 2010 (The 7th International Symposium on Chinese Spoken Language Processing)*, Tainan/Sun Moon Lake, Taiwan, pp.55-59, 2010.
- [14] S. Blum-Kulka, "Learning to Say What You Mean in a Second Language: A Study of the Speech Act Performance of Learners of Hebrew as a Second Language," *Applied Linguistics*, vol. 3 no1, pp.29-59, 1989.
- [15] Council of Europe, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, N.Y, U.S.A, Cambridge University, pp. 48-49, 2001.
- [16] Council of Europe, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, N.Y, U.S.A, Cambridge University, pp.58-61, 2001.

Appendix A: Sentences Eliciting Broad and Narrow Focus

(Broad focus)

Prompt: What happened? Answer: "Texas exported beef to Japan."

(Narrow focus)

Prompt: Which city exports beef to Japan? Answer: "TEXAS exports beef to Japan."

Prompt: What does Texas exports to Japan? Answer: "Texas exports BEEF to Japan."

Prompt: Where does Texas export beef to? Answer: "Texas exports beef to JAPAN."

Appendix B: Text of "The Cinderella Fairy Tale"

Once upon a time there was a girl called Cinderella. But everyone called her Cinders. Cinders lived with her mother and two stepsisters called Lily and Rosa. Lily and Rosa were very unfriendly and they were lazy girls. They spent all their time buying new clothes and going to parties. Poor Cinders had to wear all their old hand-me-downs! And she had to do the cleaning!

One day, a royal messenger came to announce a ball. The ball would be held at the Royal Palace, in honour of the Queen's only son, Prince William. Lily and Rosa thought

this was divine. Prince William was gorgeous, and he was looking for a bride! They dreamed of wedding bells!

When the evening of the ball arrived, Cinders had to help her sisters get ready. They were in a bad mood. They'd wanted to buy some new gowns, but their mother said that they had enough gowns. So they started shouting at Cinders. 'Find my jewels!' yelled one. 'Find my hat!' howled the other. They wanted hairbrushes, hairpins and hair spray.

When her sisters had gone, Cinders felt very down, and she cried. Suddenly, a voice said: 'Why are you crying, my dear?'. It was her fairy godmother!
The girl poured her heart out: 'Lily and Rosa have it all!' she cried, 'even though they're awful, and fat, and they're dull! And I want to go to the ball, and meet Prince William!'

'You will, won't you?' laughed her fairy godmother. 'Go into the garden and find me a pumpkin'. Cinders went, and found a splendid pumpkin which the fairy changed into a dazzling carriage.

'Now bring me four white mice,' the godmother said. The girl went, and found one... two...three...four mice. The fairy godmother changed the mice into four lovely horses to pull the carriage.

Then the girl looked at her old rags. 'Oh dear!' she sighed. 'Where will I find something to wear? I don't have a gown!' 'Hmmm...' said the fairy: 'Let's see, what do you need? You'll need a ballgown... you need jewellery... you need shoes, and... something needs to be done about your hair. And would you like a blue gown or a green gown?'

For the third time, Cinders' godmother waved her magic wand. A ballgown, a robe and jewels appeared. And there were some elegant glass slippers. 'You look wonderful,' her fairy godmother said, smiling. 'Just remember one thing - the magic only lasts until midnight!' And off Cinders went to the ball.

In the Royal Palace, everyone was amazed by the radiant girl in the beautiful ballgown. 'Who is she?' they asked. Prince William thought Cinders was the most beautiful girl he had ever seen. 'Have we met?' he asked. 'And may I have the honour of this dance?'

Prince William and Cinders danced for hours. Cinders was so glad that she failed to remember her fairy godmother's warning. Suddenly the clock chimed midnight! Cinders ran from the ballroom. 'Where are you going?' Prince William called. In her hurry, Cinders lost one of her slippers. The Prince wanted to find Cinderella, but he couldn't find the girl. 'I don't even know her name,' he sighed. But he held on to the slipper.

After the ball, the Prince was resolved to find the beauty who had stolen his heart. The glass slipper was his only clue. So he declared: 'The girl whose foot will fit this slipper shall be my wife'. And he began to search the kingdom.

Every girl in the land was willing to try on the slipper. But the slipper was always too small. When the Royal travellers arrived at Cinders' home, Lily and Rosa tried to squeeze their feet into the slipper. But it was no use; their feet were enormous! 'Do you have any other girls?' the Prince asked Cinders' mother. 'One more,' she replied. 'Oh no,' cried Lily and Rosa. 'She is much too busy!' But the Prince insisted that all girls must try the slipper.

Cinders was embarrassed. She didn't want the Prince to see her in her old apron. And her face was dirty! 'This is your daughter?' the Prince asked, amazed. But then Cinders tried on the glass slipper, and it fitted perfectly!

The Prince looked carefully at the girl's face, and he recognised her. 'It's you, my darling isn't it?' he yelled. 'Will you marry me?' Lily and Rosa were horrified. 'It was you at the ball, Cinders?' they asked. They couldn't believe it! Then Cinders married William, and they lived happily ever after.

Appendix C: An Example of the Discourse Completion Tasks

Level A2 (basic)

Situation: Introducing yourself to people you don't know or greeting people you know"

A: Hi. My name is Ping-ping Su. Call me Ping.

B: Nice to meet you, Ping. I'm Li Chang. Call me Li. Where are you from?

A: I'm from Taiwan. Are you from China?

B: Yes, that's right. I'm from Beijing. I came to Taiwan last year. I'm a sophomore now. What year are you in?

A: I'm a freshman.