English Language Teaching; Vol. 11, No. 7; 2018
ISSN 1916-4742 E-ISSN 1916-4750
Published by Canadian Center of Science and Education

Linguistic Feature Analyses of Chinese Learners of English and Contributions of Discrete Features to Perceptual Judgment

Hsueh Chu Chen[1]

[1] Department of Linguistics and Modern Language Studies, The Education University of Hong Kong, Hong Kong

Correspondence: Hsueh Chu Chen, Department of Linguistics and Modern Language Studies, The Education University of Hong Kong, Hong Kong. Tel: 852-29487376. E-mail: hsuehchu@eduhk.hk

Received: May 5, 2018 Accepted: June 4, 2018 Online Published: June 6, 2018

doi: 10.5539/elt.v11n7p22 URL: http://doi.org/10.5539/elt.v11n7p22

Abstract

This study (a) conducted a feature analysis of the spoken data of Chinese university students in pronunciation, grammar, and discourse, (b) investigated the contributions of the discrete linguistic features to the perceptual ratings on foreign accent, comprehensibility, delivery, and general language use. Ten university learners were selected from the Spoken Corpus of the English of Hong Kong and Mainland Chinese learners (http://corpus.ied.edu.hk/phonetics/), in which two speakers were paired up to conduct a five minutes interview. Three-level analyses were done to investigate Chinese learners' linguistic features. Forty listeners from four L1 language backgrounds were recruited to rate the speech samples. The results show that strongly negative correlations were found between the production and perceptual rating scores for "omission of consonant(s) in final position" "redundant article 'the'", "silent pauses" and "discourse markers," suggesting that the four features can be perceived and exert strong negative influences on perceptual judgments. Pronunciation rating had the strongest positive correlations with "foreign accentedness"; grammar rating had the strongest positive correlations with "general language use"; discourse rating had the strongest positive correlations with "general delivery", and "general language use." Regarding the rating of comprehensibility, "misuse of conjunctions" "redundant article 'the'", "silent pauses", "lengthening", and "stressing" showed strong negative correlations whereas "filled pauses" had strong positive correlations with it. Regarding the rating of foreign accentedness, strong negative correlations were found between "omission of consonant(s) in final position", "lengthening", "discourse markers", and "stressing" and the rating of "foreign accentedness".

Keywords: foreign accent, production and perception, linguistic features, corpus linguistics

1. Introduction

1.1 Learner Corpora and Language Learning

It has been suggested by many previous studies that it is of great value to analyse learner corpora as these can provide opportunities to discover the typical patterns of a certain language used by learners from real language data (e.g., Kettemann & Marko, 2002; Chen & Wang, 2016). By observing the patterns of speech data, we can spot the common features learners typically and most frequently display, which will be pedagogically helpful for students to "become aware of the features of their own inter-language and possibly stimulate a restructuring of their own language use and knowledge" (Gut, 2005, p. 1). In this light, this study conducted a corpus-based feature analysis of the spoken data of Chinese university students, including analyses of phonology, syntax, and discourse.

This paper is built upon A Spoken Corpus of the English of Hong Kong and Mainland Chinese Learners (http://corpus.ied.edu.hk/phonetics/), established by Chen et al. (2014). This corpus contains roughly 12 hours of recorded data from two speech tasks (four hours of reading aloud and eight hours of interviews) from university students in Hong Kong. The linguistic analyses of the spoken data mainly focus on the following three aspects: (1) a phonological analysis mainly focused on three areas of segmental features (vowels, consonants, and syllable structures); (2) a syntactical analysis focused on examining the most common grammatical error patterns of university students in their production of spontaneous utterances; and (3) a discourse analysis mainly focused on the patterns of silent pauses, filled pauses, lengthening, repetition, discourse markers, and stress used in speakers' spoken data. In addition, the effects of various linguistic features on listeners' perceptual judgements

on foreign accentedness, comprehensibility, language delivery, language use, and coherence were also investigated. It is hoped that this multi-faceted analysis of this spoken corpus will make valuable contributions to the field of corpora and language learning.

1.2 Perceptual Judgments on Non-native Speech

Perceptual ratings on non-native speakers' (NNSs) English speech have been widely discussed in various aspects, including foreign accent rating (e.g., Kang, 2008; Kang, Rubin, & Pickering, 2010), comprehensibility rating (e.g., Kang, 2010; Kang & Pickering, 2013; Kraut & Wulff, 2013; Munro & Derwing, 1995), and intelligibility rating (e.g., Derwing & Munro, 2005; Field, 2005; Jenkins, 2000). In the past decades, there have been a number of studies examining the various factors that may contribute to listeners' ratings of comprehensibility, intelligibility, and foreign accent. Derwing and Munro (2005) define these three terms and suggest measurement methods for these three aspects of perceptual judgment. Intelligibility is defined as the extent to which a listener actually understands an utterance; a transcription task is usually used for this. Comprehensibility is defined as a listener's perception of how difficult it is to understand an utterance, and accentedness is defined as a listener's perception of how different a speaker's accent is from that of the first language (L1) community. Both of these terms are usually measured with a scalar judgment task.

A number of studies have been focused on investigating how a range of linguistic errors (e.g., grammatical, lexical, semantic, and phonological errors) contributesto an overall intelligibility judgement by native English listeners (e.g., Derwing & Munro, 1997; Munro & Derwing, 1999). In particular, the effects of pronunciation features on speech intelligibility have been widely studied in various ways. For example, Munro and Derwing (2006) and Riney, Takada, and Ota (2000) explore the correlation between intelligibility and second language-(L2) specific segmentals; Field (2005) investigates how lexical stress placement contributes to intelligibility; Munro and Derwing (2001) are focused on the role of speech rate in perceptions of the comprehensibility and accentedness of L2 speech; Hahn (2004) studies the contribution of sentence stress on intelligibility; Kang (2010) examines the relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. These researchers have attempted to isolate the role of pronunciation, as compared to other linguistic features (e.g., grammar and discourse features), in achieving comprehensibility. The results reveal that suprasegmental features independently contribute to listeners' perceptual judgments. Accent ratings were best predicted by pitch range and word stress measures whereas comprehensibility scores were mostly associated with speech rates.

While most previous research in error analysis of L2 grammar has focused on examining the errors in written texts, relatively little has been done to discover the grammatical error patterns in L2 spoken English. Politzer and Ramirez (1973) study examines 120 Mexican-American children learning English in the United States. Based on the errors extracted from the analysis of natural speech, they propose the following error taxonomy: (1) morphology: indefinite article incorrect, possessive case incorrect, 3rdperson singular incorrect, simple past tense incorrect, past participle incorrect, and comparative adjective/adverb incorrect; (2) syntax: noun phrase, verb phrase, verb-and-verb construction, word order, and some transformations (e.g., negative transformations). McCarthy and Carter (2002) summarize the most common phenomena in spoken grammar based on their observation of frequently encountered grammatical problems in a spoken corpus: (a) indeterminate structures; (b) phrasal utterances that are communicatively complete in themselves but are not sentences; (c) aborted or incomplete structures; (d) subordinate clauses not obviously connected to any particular main clause; (e) interrupted structures with other speaker contributions intervening; (f) words whose grammatical class is unclear. Ting et al. (2010) examine the grammatical errors in the spoken English of university students who are less proficient in English. In this study, they categorize the grammatical errors into verb form, preposition, article, plurality, tense, pronoun, question, and word form to reflect the common types of errors in this study.

A recent study of Saito (2011) investigates how various linguistic factors (i.e., phonological vs. lexico-grammatical aspects of language) affect the perceptions of comprehensibility from both NS and NNS listeners' perspectives. The lexico-grammatical properties in Saito (2011) include (a) the number of words, (b) the number of clauses, (c) fluency (i.e., lexical density and the number of words per clause), (d) accuracy (i.e., the percentage of error-free clauses out of all sentences or words), and (e) complexity (i.e., the percentage of dependent clauses out of all clauses). Two conclusions can be drawn from Saito's study: (1) NNS listeners tend to be more sensitive to their own phonological errors than grammatical errors; and (2) the comprehensibility ratings of NS listeners were equally influenced by all types of linguistic errors.

A number of studies have shown clearly that discourse patterns can play a significant role in overall intelligibility. From a pedagogical standpoint, a number of empirical studies have suggested that the use of discourse

structuring cues can facilitate the comprehensibility of L2 speech. For instance, Pica et al.'s study (1990) suggests that ITAs should pay more attention to discourse structuring cues in order to enhance the comprehensibility of their spoken discourse. Tyler (1992) finds that a number of discourse factors affect listeners' ability to construct a coherent interpretation of Chinese speaker's speech, which include lexical discourse markers, patterns of repetition, prosody, anaphora (patterns of ellipsis and pronominalization, and demonstrative pronouns), syntactic incorporation (hypotactic constructions), and simple clauses (paratactic constructions). Tyler (1992) regards these discourse features as information-structuring devices that can be used as cues to enhance the comprehensibility of oral discourse. Jung (2006) explores the role of discourse markers in L2 listening comprehension and finds that discourse markers can facilitate listening comprehension by providing some guidance for meaning in spoken discourse. In his study, L2 listeners misinterpreted the text when discourse markers were missing.

Although accentedness and perceived comprehensibility are interrelated, there is no doubt that they are two separate dimensions of L2 speech (Munro & Derwing, 1999). In other words, speakers with strong foreign accents can also be rated as highly comprehensible. Understanding which aspects of foreign accents may have the strongest effects on comprehensibility is becoming more relevant for English speakers and merits more attention from language researchers. The present study was designed to explore the interrelationship among comprehensibility and foreign accentedness, with a focus on phonological features, grammatical errors, and discourse features.

Three research questions were raised to address these issues:

1) What is the relationship between production and perception of the targeted linguistic features in the phonological, grammatical, and discourse dimensions?

2) Which linguistic dimension (i.e., pronunciation, grammar, or discourse) will contribute more to the rating scores of overall judgments?

3) Which specific linguistic features produced by speakers will contribute more to the overall rating of items?

2. Method

2.1 Phase one: Linguistic Analyses

2.1.1 Participants

Ten Hong Kong speakers were randomly selected from the established spoken corpus, A Spoken Corpus of the English of Hong Kong and Mainland Chinese Learners (available at http://corpus.ied.edu.hk/phonetics/), developed by Chen et al. (2014). The speakers, aged between 18 and 22 years old, were selected from the undergraduate program of the Education University of Hong Kong. They were born in Hong Kong, and Cantonese is their L1.All of the speakers have an English proficiency level of International English Language Testing System (IELTS) 6.0 or above.

2.1.2 Targeted Linguistic Features

In the captioned corpus, two speakers were paired up to conduct a one-on-one interview. They were asked to take turns asking questions that had been prepared by the researcher. During the interview, they were also encouraged to ask further questions linked to the topic and develop their own questions based on their partners' responses. Each interview is five minutes long on average but in the current study, only two-minute recordings containing long answers were selected from each speaker for further analysis and perceptual rating. Multi-level analyses were done to investigate Hong Kong speakers' linguistic features in spoken English.

One trained coder with extensive linguistics training and experience received a detailed description for each measure, and then analysed the entire dataset. Since all phonological measures (intonation, in particular) were subject to the influence of discourse context and individual speaker variability, the coder not only used speech software, i.e., Praat (Boersma & Weenink, 2012), but also relied on her intuition as a native speaker of English for the purpose of reliable error judgements. Subsequently, another trained coder checked 40% of the speech samples. Intra-class correlations, computed to determine coding agreement, revealed high consistency values exceeding .90 for all measures. Table 2 shows the frequency table of linguistic features per 100 words.

2.2 Phase Two: Perceptual judgments

In second stage of our study, we conducted perceptual rating sessions to investigate the extent to which the features we found could be perceived. Forty listeners from four groups with various L1 language backgrounds—English as a L1 group (n=10), non-Chinese group (n=10), Hong Kong group (n=10), and Mainland Chinese group (n=10)—were recruited to rate the speech samples of the selected Hong Kong speakers

in terms of three levels of linguistic features (i.e., pronunciation, grammar, discourse). All of these participants were university students in Hong Kong and have an English proficiency level of IELTS 6.5 or above.

In addition, the contributions of the discrete linguistic features to the perceptual ratings on foreign accent, comprehensibility, delivery, and general language use were also explored. All the phonetic terms that were used in the rating sheet were introduced with concrete examples, and detailed rating criteria were explained. Appendix 1 shows the sample rating sheet.

3. Results

The following section will be structured based on the research questions addressed above. First, the results of the feature analysis will be presented. The relationship between 10 speakers' production and 40 listeners' perception of the target linguistic features at the phonological, grammatical, and discourse levels will be reported. After that, the 40 listeners' discrete rating items (i.e., phonological, grammatical, and discourse features) that contributed most to their rating scores of their overall judgments will be examined. Finally, the specific linguistic features produced by 10 speakers contributing to the overall rating items in perception will be discussed.

3.1 Feature Analysis

To answer the first research question, a feature analysis was conducted. We identified the most frequently occurring pronunciation features (n=5), grammar features (n=5), and discourse features (n=6). Table 1 presents the frequency of the linguistic features of 10 speakers that occurred per 100 words.

Table 1. Frequency table of linguistic features (per 100 words)

Features	*S1*	*S2*	*S3*	*S4*	*S5*	*S6*	*S7*	*S8*	*S9*	*S10*
Pronunciation										
Omission of consonant (s) in final position	4.0	1.8	5.6	0.0	4.8	6.7	1.5	2.0	1.9	0.0
Substitution of TH sound	2.0	1.4	1.2	0.4	2.2	1.0	3.5	4.3	1.4	0.9
Absence of contrasts between long and short vowels	0.0	2.5	0.6	0.0	2.7	0.0	0.5	1.3	2.3	0.5
Substitution of /v/→/f/	1.2	0.4	0.0	0.9	1.6	0.0	0.0	0.0	0.5	0.9
Word-level stress	0.4	0.0	0.0	0.9	0.5	1.0	0.0	0.0	0.5	0.0
Grammar										
Incorrect use of plurals	0.0	0.4	5.6	0.0	2.2	1.0	1.5	0.7	2.3	0.9
Misuse of conjunction	2.0	1.8	0.6	0.4	1.1	0.5	2.5	1.0	0.9	0.5
Redundant "the"	0.8	3.5	0.0	0.4	0.5	0.0	3.5	0.7	0.9	0.0
Incorrect verb form	0.4	2.1	1.9	0.0	0.5	0.5	0.5	1.0	0.5	0.5
Incorrect tense	0.4	1.8	3.1	0.9	1.6	0.5	1.5	2.0	0.0	1.4
Discourse										
Silent pauses	9.9	7.4	11.8	11.8	14.5	16.7	21.3	13.1	15.5	5.1
Filled pauses	9.1	2.8	18.6	18.0	5.4	7.6	4.5	6.2	9.4	11.1
Lengthening	5.6	3.5	3.7	3.9	2.7	6.2	4.0	2.6	2.8	0.5
Repetition	1.2	1.8	7.5	4.4	10.2	11.0	12.4	3.3	8.5	0.0
Discourse markers	7.1	6.3	8.7	7.5	6.5	10.0	10.4	4.3	3.8	3.2
Stressing	1.2	2.8	0.6	1.3	0.0	3.3	3.0	1.6	0.0	0.0

The top five most frequently occurring pronunciation features in our corpus are:1)omission of consonant(s) in the final position (syllable structures), 2) substitution of the "th" sound (consonants), 3) absence of contrast between long and short vowels (vowels),4)substitution of /v/→/f/ (consonants), and5)word-level stress (stress). We categorized the grammatical features into three levels: the morphological level, the lexical level, and the syntactic level. The top five most frequently occurring grammatical features are: 1) incorrect use of plurals (nouns), 2) redundancy/misuse of conjunctions (conjunctions), 3) redundant article "the" (articles), 4) incorrect verb forms (verbs), and 5) incorrect tense (verb). The discourse features were mainly analysed in terms of the following six aspects: silent pauses, filled pauses, lengthening, repetition, discourse markers, and stress. Table 2

summarizes the analysed linguistic features and examples.

Table 2. Analysed linguistic features and examples

Items	*Linguistic Features and*	*Examples*
	Phonological Features	
1	Omission of consonant (s) in final position	friend→frien, most→mos
2	Substitution of 'TH' sound	there→dere, with→wif,think→sink
3	Absence of contrasts between long and short vowels	eat→it, feel→fill
4	Substitution of /v/(/v/→/f/)	very-fery
5	Word-level stress	speCIfic→sPEcific, POlitics→poLItics
	Grammatical Features	
6	Incorrect use of plurals	many place , a lot of friend in Beijing)
7	Misuse of conjunctions	"and", "but", "even if", "because" and "so"
8	Redundant article "the"	in the Singapore, I love theHong Kong
9	Incorrect verb form	friends had went…, rules to followed…
10	Incorrect tense	I go back to (when talking about things in The past
	Discourse Features	
11	Silent pauses>0.5s	silent pauses that are longer than 0.5s
12	Filled pauses	fillers like um, uh, erm, mm, er, ah, etc.
13	Lengthening	sounds or words produced with longer duration
14	Repetition	repetitions of phrase, word, syllable and sound
15	Discourse markers	'okay', 'so', 'now', 'well', 'I mean', etc.
16	Sentential stressing	the words/ sounds stressed with pitch rise and tension

3.2 Correlation Between Production and Perception of the Target Linguistic Features

To better understand whether the 40 listeners could easily perceive the linguistic performance of the 10 speakers, Pearson's correlation coefficient (r) was computed for the features produced by the 10 speakers and the rating scores on these features from the 40listeners. Strongly negative correlations were found between the production (i.e., the number of produced features per 100 words) and the perceptual rating scores for "omission of consonant(s) in the final position," e.g.,friend→frien, most→mos (r=-0.629), "redundant article 'the,'" e.g., in the Singapore, I love the Hong Kong (r=-0.921, p<0.05), "silent pauses >0.5s," e.g., a break in speaking of 0.5 seconds or longer, which indicates hesitation (r=-0.691, p<0.05), and "discourse markers," e.g.,"well," "now," "but," "so," "then," "finally," "actually," "I mean," and "you know" (r=-0.665, p<0.05).

To give an example, discourse markers will be discussed more in-depth. Discourse markers refer to words and phrases that are used to signal the relationship and connections between utterances and to mark the beginning or end of a turn. Some common discourse markers include "first," "on the other hand," "now," "what's more," "so anyway," "well," "oh," "of course," "yeah," and "right." In the current corpus data, we notice many typical discourse markers produced by the speakers. HK4 tends to use "well" many times to mark the beginning of his turns. "Well" seems like a typical boundary for most English speakers to begin their utterances. As for the other examples from HK2, HK3, and HK5, shown in Table 3, "I think" is also frequently used by the speakers. Some possible reasons for using these discourse markers is that making the sound gives the speaker time to think about what he or she is going to say and also marks his or her beginning of the turn.It's also a way for a less-skilled speaker to hold the floor.

Table 3. Selected examples of discourse markers produced by the speakers

Speaker	*Examples*	*Discourse marker*
HK1	HK1: **actually** i am quite a lazy person (R),	actually, maybe,
	HK1: **maybe** HOLLAND (1) because i know the windmill is very attractive	
HK2	HK2: mm **i think**that this year	I think
HK3	HK3: um: (1) **i think** uh it's their (.) ar COMMUNICATIONS	I think
HK4	HK4: **well** (.) she is actually a housewife (1) yeah.	well, maybe, perhaps
	HK4: **well:** (R) in fact i: (.) seldom play ba- ba- basketball in these (.) years	
	HK4: **maybe** some kind of um (1) philosophy philosophy books	
	HK4: mm: **perhaps** (.) the united states	
HK5	HK5: **i think:** (1) many people come to hongkong (.)	I think

The results suggest that the four above-mentioned features could be perceived more easily by listeners and exerted strong negative influences on listeners' perceptual judgments. Table 4 shows the correlation between the features in production and perception.

Table 4. Correlation between 10 speakers' production and 40 listeners' perception of the target linguistic features

Effect size	***Features***	***Pearson(r)***
Large(r> 0.50, r<-0.50)	1. Omission of consonant (s) in final position	*-.629*
	8. Redundant article "the"	*-.921**
	11. Silent pauses>0.5s	*-.691**
	15. Discourse markers	*-.665**
Medium (0.3<r<0.5, -0.3>r>-0.5)	3. Absence of contrasts between long and short vowels	*-.443*
	6. Incorrect use of plurals	*-.495*
	7. Misuse of conjunctions	*-.402*
	12. Filled pauses	*-.353*
	13. Lengthening	*-.437*
	14. Repetition	*-.476*
	16. Sentential stressing	*-.389*
Small (r< 0.30, r>-0.30)	2. Substitution of 'TH' sound	*-.150*
	4. Substitution of /v/(/v/→/f/)	*-.181*
	5. Word-level stress	*-.229*
	9. Incorrect verb form	*-.208*
	10. Incorrect tense	*.015*

**. Correlation is significant at the 0.05 level (2-tailed).*

3.3 Relationship Between Three Linguistic Dimensions and Four General Impression Items

To better understand the contribution of three linguistic dimensions (e.g., pronunciation, grammar, discourse) to the general impression ratings on four of the speakers (i.e., comprehensibility, foreign accentedness, general language use, and general delivery), Pearson's correlation coefficient (r) was computed. Strongly positive correlations were found between the rating scores in pronunciation, grammar, and discourse dimensions and the

four general rating items, as shown in Table 5. Among all of the dimensions, the rating of the pronunciation dimension had the strongest positive correlation with "foreign accentedness" (r=0.944); the rating of the grammar dimension had the strongest positive correlation with "general language use" (r=0.921); the rating of the discourse dimension had the strongest positive correlation with "general delivery" (r=0.951) and "general language use" (r=0.944). Table 5 shows a correlation between rating scores in the pronunciation, grammar, discourse dimensions and the six general rating items.

Table 5. Correlation between 40 listeners' rating scores in three linguistic dimensions and the four general impression items

F		Comprehensibility	Foreign accent	Language use	Delivery
Pronunciation	r	.867	**.944**	.862	.746
	Sig.	.001	.000	.001	.013
Grammar	r	.865	.832	**.921**	.805
	Sig.	.001	.003	.000	.005
Discourse	r	.836	.843	.900	**.951**
	Sig.	.003	.002	.000	.000

3.4 Specific Linguistic Features Contributing to the Overall Impression Items

The following summarizes the specific linguistic features contributing to the rating of "comprehensibility," "foreign accentedness," "general language use," and "general delivery". Regarding the rating of comprehensibility (shown in Table 6), the results of the calculation of Pearson's correlation coefficient showed that "misuse of conjunctions" (r=-.588), "redundant article 'the'" (r=-.579), "silent pauses > 0.5s" (r=-.501), "lengthening" (r=-.668*), and "stress" (r=-.614) all had strong negative correlations with the rating of "comprehensibility." This means that the greater the occurrence in speech of the features listed above, the lower the comprehensibility ratings will be.

Table 6. Factors contributing to the rating of "comprehensibility"

Effect size	Features	Pearson correlation coefficient (r)
Large (r> 0.50, r<-0.50)	Misuse of conjunctions	-.588
	Redundant article "the"	-.579
	Silent pauses>0.5s	-.501
	Filled pauses	.528
	Lengthening	-.668*
	Stressing	-.614
Medium	Omission of consonant (s) in final position	-.365
(0.3<r<0.5, -0.3>r>-0.5)	Substitution of 'TH' sound	-.340
	Repetition	-.387
	Discourse markers	-.460
Small (r< 0.30, r>-0.30)	Absence of contrasts between long and short vowels	-.295
	Substitution of /v/(/v/→/f/)	.274
	Word-level stress	-.141
	Incorrect use of plurals	.125
	Incorrect verb form	-.266
	Incorrect tense	.145

*. Correlation is significant at the 0.05 level (2-tailed).

To give an example, we will focus on silent pauses. Riggenbach (1991) reports that the frequency of unfilled

pauses is a strong indicator of oral proficiency. Kormos and Denes (2004) maintain that features such as the number and length of pauses will affect fluency judgments to a great extent. Fillmore (1979) argues that fluency is best predicted by "the ability to talk at length with few pauses, the ability to fill time with talk" (p. 51). Lennon (2000) defines fluency as "the speed and smoothness of oral delivery" (p. 25). From the studies listed above, we can infer that the greater the occurrence of silent pauses in the utterances, the lower the speaker's fluency will be perceived. For the current speakers, silent pauses, especially those around 500ms, occurred in their narrative speeches very frequently. The exact number of silent pauses produced by the first five speakers is shown in Table 7.

Table 7. Selected examples of silent pauses produced by the speakers

Speakers	*(.)*	*(1)*	*(2)*	*Total number of silent pauses*
HK1	22	9	7	38
HK2	54	9	3	66
HK3	20	3	0	23
HK4	61	6	2	69
HK5	39	13	7	59

Notes: A half second silent pause is expressed as "(.)".

One second silent pause is expressed as "(1)".

Two seconds silent pause is expressed as "(2)".

Lengthening is another example to explore. When speakers want to emphasize specific information, they tend to make the duration of a word or a phoneme longer. Lengthening may also happen when the speakers are searching for the words that they are going to say next. By lengthening the segments, speakers can earn some time for themselves to construct in their minds the words or structures they are going to produce. However, according to the rating scores, if the sounds or words produced by a speaker were intentionally of longer duration, this reduced the listeners' comprehensibility ratings.

On the contrary, the produced "filled pauses" had a strong positive (r=.528) correlation with the comprehensibility rating. This indicates that the filled pauses in speech such as "um," "uh," "erm," "mm," "er," and "ah" will increase listeners' comprehensibility ratings. In this corpus data, the filled pause tends to occur either at the beginning of a clause to introduce a new topic or at the end of an utterance, when speakers are searching for what they are going to say next. Below are some examples extracted from the data set. The two speakers below tend to start their utterances with a filled pause (e.g., "mm,""uh"), which may earn them a little time to think of what they are going to say next.

HK1: mm i i quite like Beijing the architecture

HK1: uh if i would (.) uh i would like to go (.) mm JAPAN

HK1: ershe (2) mm: she (.) was er she is doing a part-time job

HK1: aw, Tianjin is a …

HK2: mm i would (.) want (1) i would like to go travel

Table 8 shows the factors contributing to the rating of "foreign accentedness." For the rating of foreign accentedness, strong negative correlations were found between "omission of consonant(s) in the final position" (r=-.540), "lengthening" (r=-.727), "discourse markers"(r=-.565), and "stress" (r=-.534) and the rating of "foreign accentedness."

Table 8. The factors contributing to the rating of "foreign accentedness"

Effect size	Features	Pearson correlation coefficient (r)
Large (r> 0.50, r<-0.50)	Omission of consonant (s) in final position	-.540
	Lengthening	-.727*
	Discourse markers	-.565
	Stressing	-.534
Medium (0.3<r<0.5, -0.3>r>-0.5)	Misuse of conjunctions	-.498
	Redundant article "the*"*	-.443
	Silent pauses>0.5s	-.468
	Filled pauses	.461
	Repetition	-.457
Small (r< 0.30, r>-0.30)	Substitution of 'TH' sound	-.173
	Absence of contrasts between long and short vowels	-.275
	Substitution of /v/(/v/→/f/)	.066
	Word-level stress	-.300
	Incorrect use of plurals	.063
	Incorrect verb form	-.210
	Incorrect tense	.131

*. Correlation is significant at the 0.05 level (2-tailed).

For an example, we can look at stress. Similar to lengthening, the speakers in the data put extra emphasis on the words or sounds that they want to stress by articulating them with rising pitch, longer duration, and increased volume. The speakers usually put more stress on content words that can provide the main information than they put on function words that only serve grammatical functions or connect different parts of sentences together in spoken English. Differently from native speakers of English, some of the speakers in the data tend to stress conjunctions (e.g., "so," "because," and "even though") and auxiliaries (e.g., "should").

Table 9 shows the factors contributing to the rating of "general language use". For the rating of general language use, strong negative correlations were found between "misuse of conjunctions" (r=-.657*), "redundant article 'the'"(r=-.527), "silent pauses >0.5s" (r=-.675*), "filled pauses" (r=-.529), "lengthening" (r=-.656*), "repetition" (r=-. 624), and the rating of "general language use."

Table 9. Factors contributing to the rating of "general language use"

Effect size	Features	Pearson correlation coefficient (r)
Large (r> 0.50, r<-0.50)	Misuse of conjunctions	-.657*
	Redundant article "the*"*	-.527
	Silent pauses>0.5s	-.675*
	Filled pauses	.529
	Lengthening	-.656*
	Repetition	-.624
	Discourse markers	-.578
Medium (0.3<r<0.5, -0.3>r>-0.5)	Omission of consonant (s) in final position	-.454
	Stressing	-.475
Small (r< 0.30, r>-0.30)	Substitution of 'TH' sound	-.291
	Absence of contrasts between long and short vowels	-.244
	Substitution of /v/(/v/→/f/)	.080

	Word-level stress	-.225
	Incorrect use of plurals	.012
	Incorrect verb form	-.025
	Incorrect tense	.253

*. Correlation is significant at the 0.05 level (2-tailed).

Table 10 shows the factors contributing to the rating of "general delivery." For this rating, strong negative correlations have been found between "misuse of conjunctions" (r=-.701*), "redundant article 'the'" (r=-.555), "silent pauses >0.5s" (r=-.654*), "lengthening" (r=-.682), "repetition" (r=-.524) and "discourse markers" (r=-.593) and the rating of "general language use."

Table 10. Factors contributing to the rating of "general delivery"

Effect size	Features	Pearson correlation coefficient (r)
Large (r> 0.50, r<-0.50)	Misuse of conjunctions	-.701*
	Redundant article "the"	-.555
	Silent pauses>0.5s	-.654*
	Lengthening	-.682
	Repetition	-.524
	Discourse markers	-.593
Medium (0.3<r<0.5, -0.3>r>-0.5)	Incorrect tense	.324
	Filled pauses	.418
	Stressing	-.481
Small (r< 0.30, r>-0.30)	Omission of consonant (s) in final position	-.263
	Substitution of 'TH' sound	-.280
	Absence of contrasts between long and short vowels	-.100
	Substitution of /v/(/v/→/f/)	.001
	Word-level stress	-.286
	Incorrect use of plurals	.171
	Incorrect verb form	.126

*. Correlation is significant at the 0.05 level (2-tailed).

It is worth noting that "substitution of /v/ (/v/→/f/)," "incorrect use of plurals," and "incorrect verb form" had a small effect on all four rating items; in other words, the occurrence of these three features did not have any significant effect on the rating of "comprehensibility," "foreign accentedness," "general language use," and "general delivery." Interestingly, we also found some linguistic features that exerted little influence on the overall perceptual judgments; for example, "incorrect tense" had very little correlation with perceived "comprehensibility" (r=.145) and "foreign accentedness" (r=.131), "substitution of the 'th' sound" had very little correlation with "foreign accentedness" (r=-.173), and "word-level stress" had very little correlation with perceived "comprehensibility"(r=-.141).

4. Discussion and Conclusion

The present study investigated in detail how people perceive phonological, grammatical, and discourse aspects of language in accentedness, comprehensibility, and general language use and delivery. In most Chinese English as a second language (ESL) classrooms, it is observed that a great deal of attention is directed at grammar instruction, but the value of teaching pronunciation and discourse has not been sufficiently recognized by researchers, teachers, and learners. In order to meet L2 learners' need for effective learning materials, it is time for L2 teachers to approach second language acquisition processes not only in the syntactic domain but also in the phonological and discourse domains in a more interdisciplinary manner. This section presents pedagogical issues regarding pronunciation and discourse teaching in L2 classrooms.

This study found that there are strong influences on the rating of foreign accentedness from "omission of consonant(s) in the final position," "lengthening," "discourse markers," and "stress." "Filled pauses," "silent pauses," "lengthening," and "stress" had a strong impact on the comprehensibility ratings. Based on these findings, improving comprehensibility and reducing accentedness would most likely involve an integrative approach targeting the crucial pronunciation features of pausing, stress, lengthening, and discourse markers that affect successful L2 communication. For instance, teaching Chinese learners to achieve beginner-level comprehensibility would include the development of optimal fluency (Munro & Derwing, 2001) and appropriate prosody (Field, 2005). Burgess and Spencer (2000) propose the second- or foreign-language teaching and learning model, which attempts to incorporate pronunciation into integrated skills through listening tasks or reading discourse or text. This model enhances students' pronunciation skills through integrated practice within two areas: the general communicative framework (e.g., listening and speaking) and communicative language teaching methods concerning other aspects of language (e.g., discourse features). The model focuses on practicing pronunciation in three-stage activities, namely input 1, output 1, and input/output 2, so that students' articulation of certain language features can be enhanced.

At the input 1 stage, for example, a language focus such as word-final consonant clusters is introduced through listening to or reading a target text. All the words that include this feature are highlighted. The teacher plays an excerpt of a native speaker or reads aloud a passage with the target feature. This is an effective way to provide learners with explicit vocabulary instruction, particularly targeting words with the accurate use of consonants in the final position in L2 oral discourse while at the same time helping students pronounce these words at an optimal speaking rate (e.g., pausing, lengthening) and with proper prosody (e.g., stress).

At the output 1 stage, ideational frameworks can be used to facilitate the processing of ideational content (Burgess & Carter, 1996). Flow diagrams, tree diagrams, or grids can serve as ideational frameworks, which can be the basis for the controlled speaking work at this stage. When the students are asked some questions, they can answer them with the assistance of the flow diagrams. The language focus (consonant clusters in the final position) are highlighted.

At the input/output 2 stage, the emphasis is placed on what the students are communicating while how they are communicating it is not the main focus. During these stages of L2 comprehensibility development, teachers might wish to encourage learners to produce different types of words through various kinds of meaning-focused input and output tasks while at the same time drawing their attention to segmental errors during such tasks via a range of interactive feedback techniques (Saito, 2013). Various activities can be developed with a stronger focus on the interpersonal information expressed in the input 1 stage and fluency in speaking, with the purpose of allowing learners to make use of the target language outside the classroom context (Burgess & Spencer, 2000). The teacher, for instance, can invite the students to express their opinions through discussion, such as "What do you think is the moral of the story?" with emphasis on using the language (e.g., discourse markers, stress) for interpersonal purposes. In order to enhance students' phonological and prosodic awareness, word banks that contain vocabulary with consonant clusters in the final position should be introduced to students to enhance their vocabulary range with the phonological feature.

In summary, this study focused on two factors derived from listeners—comprehensibility and accentedness—by examining how different aspects of language contribute to these factors. We as teachers should guide students in deciding whether they want to pursue ease of understanding or producing native like speech as their learning goal. In different social, academic, or business settings, learners may have different learning goals to achieve. It is essential to set achievable goals for ESL adult learners by prioritizing understanding or native likeness in order to be able to communicate successfully.

References

Albrechtsen, D., Henriksen, B., & Færch, C. (1980). Native speaker reactions to learners' spoken interlanguage. *Language Learning, 30,* 365-396. https://doi.org/10.1111/j.1467-1770.1980.tb00324.x

Anderson-Hsieh, J., & Koehler, K. (1988).The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning, 38,* 561-613. https://doi.org/10.1111/j.1467-1770.1988.tb00167.x

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992).The relationship between native speaker judgements of nonnative pronunciation and deviance in segmentals, prosody and syllable structure.*Language Learning, 42*, 529-555. https://doi.org/10.1111/j.1467-1770.1992.tb01043.x

Boersma, P., &Weenink, D. (2012). *Praat: doing phonetics by computer* [Computer program].

Burgess, J., & Carter, I.G. (1996). Common codes for mainstream ESL support across the curriculum. *System,*

24(2), 211-222. https://doi.org/10.1016/0346-251X(96)85377-2

Burgess, J., & Spencer, S. (2000). *Phonology and pronunciation in integrated language teaching and teacher education. System, 28*, 191-215. https://doi.org/10.1016/S0346-251X(00)00007-5

Chen, H. C., Wang, L., Wong, P. M. J., & Chan, K. Y. (2014). The Spoken Corpus of the English of Hong Kong and Mainland Chinese learners. The Hong Kong Institute of Education. Retrieved from http://corpus.ied.edu.hk/phonetics/

Chen, H. C., & Wang, Q. (2016). Development and application of a corpus-based online pronunciation learning system for Chinese learners of English. *English Teaching & Learning, 40*(2), 83-115.

Derwing, T. (1990). Speech rate is no simple matter: Rate adjustment and NS-NNS communicativesuccess. *Studies in Second Language Acquisition, 12*, 303-313. https://doi.org/10.1017/S0272263100009189

Derwing, T., & Munro, M. (1997).Accent, intelligibility, and comprehensibility.*Studies in Second Language Acquisition, 12*, 1-16. https://doi.org/10.1017/S0272263197001010

Derwing, T., & Munro, M. (2005). Second language accent and pronunciation teaching: A research basedapproach. *TESOL Quarterly, 39*, 379-397. https://doi.org/10.2307/3588486

Ensz, K.Y. (1982). French attitudes toward typical speech errors of American speakers of French. *The Modern Language Journal, 66*, 133-139. https://doi.org/10.1111/j.1540-4781.1982.tb06972.x

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly, 39*(3), 399-423. https://doi.org/10.2307/3588487

Fillmore, C. J. (1979). On fluency. In H. Riggenbach (Eds.), *Perspectives on Fluency*. Ann Arbor, MI: University of Michigan Press. https://doi.org/10.1016/B978-0-12-255950-1.50012-3

Flowerdew, J. (1994). *Academic listening: Research perspectives*. New York: Cambridge University Press.

Gut, U. (2005). The realisation of final plosives in Singapore English: Phonological rules and ethnic differences. In D. Deterding, A. Brown, & E. L. Low (Eds.), English in Singapore: Phonetic research on a Corpus (pp. 14-25). Singapore: McGraw-Hill.

Hahn, L. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38,* 201-223. https://doi.org/10.2307/3588378

Hatch, E. (1992).*Discourse and language education.*Cambridge: Cambridge University Press

Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition, 34*(3), 475-505. https://doi.org/10.1017/S0272263112000150

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford UniversityPress.

Jesney, K. (2004). The use of global foreign accent rating in studies of L2 acquisition. Calgary, AB: Language Research Centre, University of Calgary.

Jung, E. H. (2006). Misunderstanding of academic monologues by non-native speakers of English.*Journal of Pragmatics, 38*(11), 1928-1942. https://doi.org/10.1016/j.pragma.2005.05.001

Kang, O., & Pickering, L. (2013). Using acoustic and temporal analysis for assessing speaking. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1047-062). Hoboken, NJ: Wiley-Blackwell. https://doi.org/10.1002/9781118411360.wbcla056

Kang, O. (2008). The effect of rater background characteristics on the rating of International Teaching Assistants Speaking Proficiency. *Spaan Fellow Working Papers*, *6*, 181-205.

Kang, O. (2010). Relative Salience of Suprasegmental Features on Judgments of L2 Comprehensibility and Accentedness.*System: An International Journal of Educational Technology and Applied Linguistics, 38*(2), 301-315. https://doi.org/10.1016/j.system.2010.01.005

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal, 94*, 554-566. https://doi.org/10.1111/j.1540-4781.2010.01091.x

Kormos, J., & Denes, M. (2004). Exploring measures and perceptions of fluency in the speech.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Eds.), *Perspectives on Fluency*. Ann Arbor, MI: University of Michigan Press.

Major, B. N., Eliezer, D., & Rieck, H. (2012). The psychological weight of weight stigma. *Social Psychological and Personality Science, 3*, 651-658. https://doi.org/10.1177/1948550611434400

Munro, M. J., & Derwing, T. M. (1995). Processing time, accent and comprehensibility in the perception ofnative and foreign accented speech. *Language and Speech, 38*, 289-306. https://doi.org/10.1177/002383099503800305

Munro, M. J., & Derwing, T. M. (2001). Modeling Perceptions of the Accentedness and Comprehensibilityof L2 Speech: The Role of Speaking Rate. *Studies in Second Language Acquisition, 2*3(4), 451-468.

Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The Mutual Intelligibility of L2 Speech. *Studies in Second Language Acquisition, 28*(1), 111-131. https://doi.org/10.1017/S0272263106060049

Munro, M. J., & Derwing, T. (2006). The functional load principle in ESL pronunciation instruction:An exploratory study. *System, 34*, 520-531. https://doi.org/10.1016/j.system.2006.09.004

Munro, M. J., & Derwing, T. (1999). Foreign accent, comprehensibility, and intelligibility in the speechof second language learners.*Language Learning, 49*(Suppl. 1), 285-310. https://doi.org/10.1111/0023-8333.49.s1.8

Pica, T., Barnes, G., & Finger, A. (1990). *Teaching matters: Skills and strategies for international teaching assistants.* New York: Newbury House.

Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*, 191-215. https://doi.org/10.1006/jpho.2001.0134

Politzer, R., & Ramirez, A. (1973). An error analysis of the spoken English of Mexican-American pupils in a bilinguag school and monolingual school. *Language Learning, 23*(1). https://doi.org/10.1111/j.1467-1770.1973.tb00096.x

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of non-native speaker conversations. Discourse Process. https://doi.org/10.1080/01638539109544795

Riney, T., Takada, M., & Ota, M. (2000). Segmentals and global foreign accent: The Japanese flap in EFL. *TESOL Quarterly, 34*, 711-737. https://doi.org/10.2307/3587782

Saito, K. (2013c). Re-examining effects of form-focused instruction on L2 pronunciation development: The role of explicit phonetic information.*Studies in Second Language Acquisition, 35*, 1-29. https://doi.org/10.1017/S0272263112000666

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research 12*(3), 329-363. https://doi.org/10.1177/1362168808089921

Tyler, A, (1992). Discourse structure and specification of relationships: A cross- linguistic analysis. *Interdiciplinary Journal for the Study of Discourse, 12,* 1-18. https://doi.org/10.1515/text.1.1992.12.1.1

Tyler, A, (1994). The role of syntactic structure in discourse structure: signaling logical and prominence relations. *Applied Linguistics, 15*, 243-262. https://doi.org/10.1093/applin/15.3.243

Yeni-Komshian, G., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean-English bilinguals. *Bilingualism: Language and Cognition, 3*(2), 31-149. https://doi.org/10.1017/S1366728900000225

Appendix A

Sample rating sheet

Instruction: Listen to the recording and rate the interviewee's speech performance in following 3 aspects (i.e. pronunciation, grammar, and discourse features). After each sentence, there is a 5-second pause so that you can mark your answers by circling the number in the bipolar rating scale. A score of 1 is the most negative, and 5 is the most positive. At the end of each recording, you will be given 1 minute to complete your rating in both Part 1 and Part 2.

Part 1. Perception of linguistic features in the speech

Pronunciation features						
		very serious	**1**	**2**	**3**	**4** **5** **not serious**
1.	***Omission of consonant (s) in final position*** *(e.g. friend→frien, most→mos)*	1	2	3	4	5
2.	***Substitution of 'TH' sound (****e.g. there→dere, with→wif, think→sink)*	1	2	3	4	5
3.	***Absence of contrasts between long and short vowels (****e.g. eat→it, feel→fill)*	1	2	3	4	5
4.	***Substitution of /v/(/w/→/f/)*** *(e.g. very-fery)*	1	2	3	4	5
5.	***Word-level stress (****e.g.speCIfic→sPEcific, POlitics→poLItics)*	1	2	3	4	5
Grammar features						
		very serious	**1**	**2**	**3**	**4** **5** **not serious**
1.	***Incorrect use of plurals*** *(e.g.many place , a lot of friend in Beijing)*	1	2	3	4	5
2.	***Misuse of conjunctions (****e.g. "and", "but", "even if", "because" and "so")*	1	2	3	4	5
3.	***Redundant article "the" (****e.g.in the Singapore, I love the Hong Kong)*	1	2	3	4	5
4.	***Incorrect verb form (****e.g. friends had went…, rules to followed…)*	1	2	3	4	5
5.	***Incorrect tense*** *(e.g. I go back to (when talking about things in the past)*	1	2	3	4	5
Discourse features						
		least appropriate	**1**	**2**	**3**	**4** **5** **most appropriate**
1.	***Silent pauses>0.5s*** *(i.e. silent pauses that are longer than 0.5s)*	1	2	3	4	5
2.	***Filled pauses*** *(i.e. fillers like um, uh, erm, mm, er, ah, etc.)*	1	2	3	4	5
3.	***Lengthening*** *(i.e. sounds or words that are produced with longer duration)*	1	2	3	4	5
4.	***Repetition*** *(i.e. repetitions of phrase, word, syllable and sound)*	1	2	3	4	5
5.	***Discourse markers*** *(i.e. 'okay', 'so', 'now', 'well', 'I mean', etc.)*	1	2	3	4	5
6.	***Stressing*** *(i.e. the words/ sounds that are stressed with pitch rise and tension)*	1	2	3	4	5

Part 2.General impression of the speaker

Incomprehensible	**1**	**2**	**3**	**4**	**5**	**Highly comprehensible**
Strong foreign accent: definitely non-native	**1**	**2**	**3**	**4**	**5**	**No foreign accent at all: definitely native**
Bad language use (bad use of grammar and vocabulary)	**1**	**2**	**3**	**4**	**5**	**Good language use** (good use of grammar and vocabulary)
Bad delivery (choppy, fragmented, or telegraphic delivery; frequent pauses and hesitations.)	**1**	**2**	**3**	**4**	**5**	**Good delivery** (fluid expression with well-paced flow , clear, and proper intonation patterns)
Bad pronunciation (uses a limited range of pronunciation features and speech is often unintelligible)	**1**	**2**	**3**	**4**	**5**	**Good pronunciation** (uses a full range of pronunciation features with precision and subtlety)
Incoherent (cannot use proper connectives and discourse markers to link simple sentences)	**1**	**2**	**3**	**4**	**5**	**Coherent** (speaks coherently with fully appropriate cohesive features and discourse markers)