System 96 (2021) 102392

Contents lists available at ScienceDirect

System

journal homepage: www.elsevier.com/locate/system

A study of Chinese EFL learners' problems with stress acquisition

Youhong Bu [a, *], Zhengdong Zhou [b]

[a] Foreign Languages College, Shanghai Normal University, No.100 Guilin Rd, Shanghai, 200234, China
[b] College of Foreign Languages and Literature, Fudan University, No. 220 Handan Rd, Shanghai, 200433, China

ARTICLE INFO

Article history:
Received 4 June 2020
Received in revised form 13 October 2020
Accepted 21 October 2020
Available online 28 October 2020

Keywords:
English stress
Intensity
Pitch
Duration
Chinese EFL learners' problems

ABSTRACT

This paper reports on a study of Chinese EFL learners' problems with the acquisition of English stress. The data were collected from 25 participants taking part in three recording tests designed by the authors. The three tests included: (1) a word list without any stress placement indicators or phonetic symbols, (2) a word list with stress placement indicators, and (3) three sentences with intonation phrase boundary marks. The results show that assigning wrong lexical stress, making the unstressed syllable almost as prominent as the stressed syllable in words, and giving inappropriate prominence to syllables in sentences are the significant factors contributing to a Chinese-English accent. Compared with previous studies on the L2 acquisition of English stress, this research offers a more detailed description of Chinese speakers' problems which have not been fully discussed so far. In addition, the authors analyze major reasons for the students' low performance, and suggest some pedagogical implications to deal with the errors detected. These implications, together with some applicable practice, will help develop the learners' oral competence and prevent interference errors from being fossilized.

1. Introduction

Much of the recent research on the acquisition of pronunciation among Chinese EFL (English as a Foreign Language) learners has switched from segmental phonemes to suprasegmental features, and particularly to intonation (Bi & Chen, 2013; Chen, 2008; Gao et al., 2015; Meng & Wang, 2009; Wang, 2011). However, few studies have explored the acquisition of English stress. In fact, stress is a fundamental element of English intonation upon which the pitch contour is formed; in other words, appropriate pitch patterns depend on the correct articulation of word stress and sentence stress.

Lewis and Deterding (2018), Lewis (2017), Guo and Chen (2017), Yang (2012), Zhang (2008), Field (2005), and Chen et al. (2001) have all investigated stress at lexical and sentential level, but their discussions mostly involve intelligibility and comprehensibility from the perspective of oral communication. The stimuli they chose for word stress were mainly disyllabic words, such as nouns versus verbs *convert*, *conduct*, *subject*, *object*, *project*, *record* (Lewis & Deterding, 2018), while those for

* Corresponding author. College of Foreign Languages and Literature, Fudan University, No. 220 Handan Rd, Shanghai, 200433, China.
E-mail addresses: lnp3yb@163.com (Y. Bu), zzd463@163.com (Z. Zhou).

https://doi.org/10.1016/j.system.2020.102392
0346-251X/

sentence stress were of only one type, concentrating on different information focuses, such as "***I*** bought a cat there," "I ***bought*** a cat there," "I bought a ***cat*** there," and "I bought a cat ***there***" (Chen et al., 2001). However, these issues alone cannot reveal Chinese EFL learners' overall problems with English stress acquisition, because their difficulties are far greater than those.

Having analyzed the stress production of participants and compared their syllable duration, pitch prominence, and vowel reduction with those of a native speaker, we found that the Chinese English-speakers showed less differentiation in syllable duration, pitch prominence, and vowel reduction than the native speaker did. The EFL learners have difficulty in distinguishing the stressed syllable from the unstressed syllable and the weakly stressed syllable of the function words. They tend to: 1) assign incorrect stress placement to polysyllabic words, particularly derivational and compound words; 2) make the unstressed syllable almost as prominent as the stressed syllable in words; and 3) give inappropriate prominence to syllables in sentences and show no information focus on the nuclear stressed syllable.

This discovery reveals that the differences between Chinese and English sound patterns contribute to Chinese EFL learners' problems with stress acquisition. As "the rhythmic basis of the two languages relies on the stress and tone" (Xu, 2013, p. xvii), stress is to English what tone is to Chinese. In other words, the phonological type of English is stress-oriented. Despite the fact that it has "tone" (pitch), i.e., pitch variations, this actually refers to the domain of analysis at the level of intonation, which is conceptually different from the monosyllabic tones in Chinese. In contrast, tone is the phonological element in Chinese. Although it also has "stress", this specifically indicates the projection and control of discourse stress on the word or phrase layer, which is distinctly different from the word stress in English. This distinction is reinforced by Xu and Shen (2016, p. 645) and Kijak (2009, p. 157), who note that Chinese stress is not operable at the lexical layer, in either the contrast layer or the boundary layer, and it has no effect on word recognition.

The critical finding of this research is that the learning of English intonation should be based, first of all, on stress. It is only by acquiring the correct word and sentence stress that the proper pitch patterns of English intonation will be developed. This is particularly important for Chinese EFL learners, whose native language is tone-oriented and gives rise to "stress-deafness" (Peperkamp & Dupoux, 2002). Thus, only through a program of systemic and reinforcement training can Chinese EFL learners overcome the problem of "stress-deafness" and acquire native-like pronunciation.

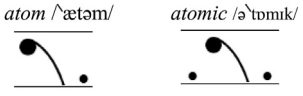
2. Acoustic properties of English stress

The differences between stressed and unstressed syllables are much greater in English, while in the Chinese language there are no such differences at all because each syllable is spoken with approximately the same amount of time and prominence, i.e., one syllable, one word, plus one tone. Thus, each full-tone syllable has nearly the same degree of loudness and duration. However, English is a stress-timed language; certain syllables are louder and longer, while others are shorter and softer and have reduced vowel quality.

2.1. Characteristics of stressed and unstressed syllables of a word

According to Ladefoged and Johnson (2015), Celce-Murcia et al. (2010, 1996), Roach (2009), Cruttenden (2008), and Ladefoged (2003), stressed syllables bear four physiological and acoustic properties: intensity/loudness, vowel duration, pitch height, and vowel quality. On the other hand, unstressed syllables lack all these features and tend to have more centralized or neutralized weak vowels /ə//i/or/u/.[1] Bolinger (1958) states that "the syllable on which the stress falls when a given word is uttered in isolation is said to have 'inherent' or 'potential' stress, or simply the word stress."

To illustrate this point more clearly, we can take the word *atom* and its derivation form *atomic* for example. The first stressed syllable of *atom* /ˈætəm/ (● •) and the second stressed syllable of *atomic* /əˈtɒmɪk/ (• ● •) are articulated with greater breath effort and muscular energy than the unstressed syllables. They are perceived to be louder, longer, and higher in pitch. Also, in the stressed syllable both vowels have a full vowel quality, as /æ/ in *atom* and /ɒ/ in *atomic*, whereas in the unstressed syllables, the vowels are reduced to either /ə/ or /ɪ/, and are perceived by the listener to be weaker and shorter than their stressed counterpart. The stressed syllable of a word is articulated with greater amplitude, longer duration, and higher pitch movement. In the case of a fall, the pitch moves from high to low, thus:



In fact, a number of studies have explored the acoustic correlates of lexical stress in English (Beckman, 1986; Bolinger, 1958; Campbell & Beckman, 1997; Cutler, 2015; Fry, 1955, 1958, 1965; Lieberman, 1960, 1975).

[1] See Wells' (2008, p. 539) *Longman Pronunciation Dictionary* about the reduced vowels/i/and/u/.

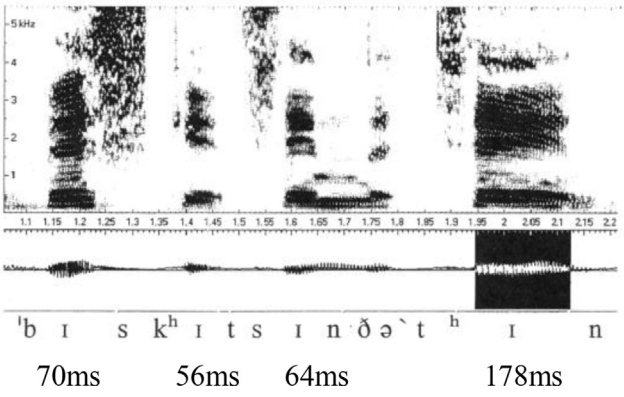
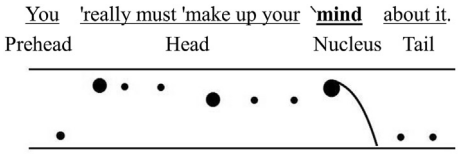


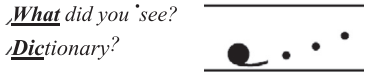
Fig. 1. The duration of the vowel/ɪ/in the phrase *biscuits in the tin*.

2.2. Characteristics of a sentence stress

Wells (2006) regards stressed syllables as those which take accent, and thus as nucleus tones in contrast to all other syllables which are pronounced with equal or nearly equal force. Wells' nucleus is the nearest to the pitch height and prominence that we describe in this article. The principal difference is that the nucleus is the major sentence stress, and other syllabic prominences due to pitch are "head". For example[2]:



The heads start from the onset, the first stress word *really*, and the second stressed word *make*, to the last word *your*, forming a series of level steps, each one lower pitched than the preceding one. Any unstressed syllables are at the same pitch height as the stressed syllable they follow. For this reason, this is sometimes known as a "stepping head" (see also Cruttenden, 2008; O'Connor & Arnold, 1973; Roach, 2009; Tench, 1996). There are as many steps as there are stressed syllables in the head. The prehead (*you*), which consists of unstressed syllable(s), is usually pitched low, while the nucleus (*mind*) is pitched the highest, as can be seen in the example with the falling tone. The syllables occurring after the nuclear syllable (*about it*) are the tail, and are pitched as low as the prehead in a falling nuclear tone, but gradually rise high in a rising nuclear tone. For example:



The pitch movement of the nuclear syllable *what* and *dic*-begins low, and is completed over the rest of the intonation phrase, i.e., the tail, *did you see* and *-tionary*.

Apart from pitch variations, the duration of the stressed syllable also varies a great deal from the unstressed syllable. Heselwood (2005) states that length differences are not random, and that the duration of the same vowel may differ in different phonetic contexts — for example, the vowel/ɪ/in the phrase *biscuits in the tin* (Fig. 1).

The longest duration is *tin* (178 ms), which is nearly three times longer than the other three syllables because it is the nuclear syllable. The vowel/ɪ/in *-kits* has the shortest duration because it appears in an unstressed syllable and before the voiceless consonant/t/. Although the word *in* is weakly-stressed, it comes before the voiced consonant/n/, and thus is a little longer (64 ms) than *-kits* (56 ms). The vowel/ɪ/in the syllable *bis-*is the next longest (70 ms) as it appears in the stressed syllable. Therefore, the duration is determined by the voice feature of following phoneme and by prosodic structure (stress and tonicity).

Therefore, the weakly-stressed syllable of function words is also a significant factor in sentence stress. These are usually distinguished by a change in vowel quality from a border position on the vowel quadrilateral to a central position (Heselwood,

[2] The example is adapted from Wells (2006, p. 212) to present a complete component of an intonation phrase and pitch patterns of a sentence.

2005). The vowel in a weak form is usually the schwa (/ə/), and next comes/i/or/u/. Weak forms are pronounced more quickly and at lower volume in comparison to strong forms (Johnson, 2012; Ladefoged, 2003; Roach, 2009).

Moreover, sentence stress is concerned with the stress placed on words in order to contrast, to emphasize, or to give new information to indicate prominence in relation to other words in a sentence. The nuclear syllable is normally placed on the last content word of an utterance to convey either a statement or a question. However, the nuclear stress can also be placed on virtually any syllable within a sentence (i.e., sentence stress), in which case the nuclear stress has a contrastive value (as in *He doesn't live* ***\in*** *Auckland* [but nearby]) (Clark & Yallop, 1995).

In short, the degree of the stressed syllable, the nuclear syllable, the unstressed syllable, and the weakly stressed syllable of function words is indicated by changes in pitch prominence and vowel duration (Berinstein, 1979; Crystal, 1969; Potisuk et al., 1996). Differences between English and Chinese sentence stress have been reviewed by Cooper and Sorensen (1981), Chao (1968), Ho (1976), and Pike (1948).

3. Methods

3.1. Participants

The participants were 25 English majors in the class of 2018 at the Foreign Languages College, Shanghai Normal University. They were all taking *English Pronunciation* in the first semester and *English Intonation* in the second semester as required courses. Their ages were between 19 and 20 at the time of recording. They came from different parts of China, and included 1 from Tibet and 2 from Ürümqi. Problems caused by their regional accents are not our concern here; we are focusing on common mistakes in the acquisition of English stress. Therefore, local accent is not a significant factor in this study.

The anonymous native speaker (NS) is a well-trained female phonetician who speaks Received Pronunciation (RP).

3.2. Materials

The data were taken from the participants' three weekly recording tests designed by the authors. The three tests included: (1) a word list without any stress placement indicators or phonetic symbols, (2) a word list with stress placement indicators, and (3) three sentences with a single boundary mark (|) for a short pause and a double boundary line (‖) for a long pause.

3.3. Instruments

In this study, Praat V6.0.55 (Boersma & Weenink, 2019) was used for recording speech, annotating audio files, extracting acoustic data, and visualizing acoustic information. SPSS 24.0 was also applied for the statistical analysis of acoustic data.

4. Results and discussion

4.1. Misplacement of lexical stress

The Chinese learners had considerable difficulty with the English lexical stress assignment, partly due to the complexity and irregularity of the English lexical stress placement, and partly because of the lack of the similar stress patterns between the two languages. "The key to the difference is that polysyllabic words are the domain of English stress assignment and monosyllabic words are the domain of Chinese tone assignment" (Xu, 2019, p. 54). The test result is consistent with Xu's remarks, and shows that the subjects were unaware of the significance of English stress, and therefore the incorrect assignment of lexical stress occurs frequently.

113 test words are arranged in accordance with their syllable types, such as disyllabic words with the distinctive feature of noun vs. verb, polysyllabic words, and words with suffixation morphemes. The stress of compounds is also taken into consideration.

4.1.1. Incorrect assignment in words with disyllables

A number of English words, the majority of which are disyllabic, are distinguished by the placement of the stress. As nouns (or adjectives) they have the stress on the first syllable, as verbs on the second — e.g. '*present* (n. or adj.) vs. *pre*'*sent* (v.). These pairs are confusing to the CLs because they are not used to making such a distinction in the Chinese language. Important as the distinctive value of the stress is, the CLs do not take them seriously. In the 20 pairs (see Table 1), the nouns '*conduct*, '*contrast*, '*digest*, and '*discount*, for instance, were pronounced as the verbs 'con'duct', 'con'trast', 'di'gest' and 'dis'count'. The verbs *ab*'*sent*, *per*'*fect*, and *pre*'*sent*, on the other hand, were confused with the nouns or adjectives of the same spelling, and pronounced as ''absent', ''perfect', and ''present'. For example, in the proverb '*practice makes perfect*', the adjective *perfect* is read as 'per'fect'.

Except for the three pairs '*object* (n.) vs. *ob*'*ject* (v.), '*record* (n.) vs. *re*'*cord* (v.), and '*subject* (n.) vs. *sub*'*ject* (v.), the other seventeen pairs were more or less mispronounced. The subjects were unable to distinguish nouns from verbs, or verbs from nouns. This result is surprising, but it is clearly indicated here.

Table 1
Incorrect assignment in words with disyllables (1000 tokens).

Nouns or Adjectives	Verbs	Nouns or Adjectives	Verbs
accent	accent	object	object
abstract	abstract	perfect	perfect
absent	absent	permit	permit
conduct	conduct	produce	produce
contrast	contrast	progress	progress
digest	digest	protest	protest
discount	discount	record	record
increase	increase	refuse	refuse
insult	insult	subject	subject
present	present	transfer	transfer

4.1.2. Polysyllabic words that tend to be assigned the wrong stress

This type of word contains different parts of speech: nouns, verbs, adjectives, and adverbs. As they have a high rate of occurrence in daily conversations, incorrect pronunciations are easily identified and seen as a sign of low proficiency in English speech. Originally, the order of the word list was not arranged as it is in Table 2; rather, the words were presented at random. However, during the process of data analysis we found that the mispronounced words followed certain regularities, and we categorized and arranged them into three groups (see Table 2): (1) words ending with a diphthong, (2) words ending with a long vowel, and (3) words with prefix plus root.

The Chinese learners tend to shift the stress to the final diphthong syllable, as can be seen in the words ending with diphthongs. Thus, they mispronounce *classify* [klæsɪ'faɪ] for /'klæsɪfaɪ/, *detail* [di:'teɪl] for /'di:teɪl/, *fertile* [fɜ:'taɪl] for/'fɜ:taɪl/, and *imitate* [ɪmɪ'teɪt] for/'ɪmɪteɪt/. The word *marmalade*/'mɑ:məleɪd/is wrongly pronounced as [ˌmɑ:mə'leɪd], since they may assume that it has the same stress pattern as the word *lemonade* (/ˌlemə'neɪd/).

Similarly, words ending with a long unstressed vowel, such as *absolute*/'æbsəlu:t/, *contribute*/'kɒntrɪbju:t/(or/kən'trɪbju:t/), *concrete*/'kɒŋkri:t/, *obsolete*/'ɒbsəli:t/, and *substitute*/'sʌbstɪtju:t/, are mispronounced as [æbsə'lu:t], [kəntrɪ'bju:t], [kɒŋ'kri:t], [ɒbsə'li:t], and [sʌbstɪ'tju:t] respectively. As with words ending with diphthongs, the participants are inclined to put the stress on the final long vowels. In fact, concerning these wrongly pronounced words, we have to admit that the uncertainty may have been caused by the inadequacies of English orthography.

The mispronounced words *agriculture* [ˌægrɪ'kʌltʃə] for/'ægrɪˌkʌltʃə/, *infamous* [ɪn'feɪməs] for/'ɪnfəməs/, *interview* [ɪntə'vju:] for/'ɪntəvju:/, *triangle* [traɪ'æŋgl] for/'traɪæŋgl/, and *vegetable* [vedʒ'teɪbəl] for/'vedʒtəbl/can be explained by the participants assuming that they sound like the single words they know, and therefore putting the stress on the root words *culture*, *famous*, *view*, *angle*, and table.

What is concerning was the high percentage of wrongly pronounced words. Words with prefix plus root show the highest occurrence of mispronunciation. If the students are accustomed to saying *infamous* [ɪn'feɪməs], for example, how can they understand the word/'ɪnfəməs/when it is said correctly? The stressed syllable is the first (not the second), and the second syllable is not a strongly stressed vowel/eɪ/but a reduced vowel/ə/. The change from/eɪ/to/ə/means that one may find it hard to recognize the word one expects to hear. Consequently, misunderstanding resulting from the incorrect placement of stress will inevitably interfere with listening comprehension.

Among the words ending with a long vowel, *absolute* and *substitute* are the two that are used most frequently in daily communication. Although the mispronunciation may not hinder mutual understanding, nevertheless it will make the speaker sound odd to their listener or audience. *Concrete* and *contribute* are also commonly used words, but nobody pronounced them correctly. The results of the survey are completely consistent with our observations of the students' actual performance in spoken English.

Words ending with a diphthong have a relatively low occurrence of errors. The reason for this is that *classify*, *detail*, *fertile*, and *imitate* happen so frequently and the correct pronunciation takes such a predominant position that the subjects may have gained some awareness of them and made self-corrections. The only exception is the word *marmalade* because few CLs get this right; even advanced learners make mistakes.

Table 2
Words that tend to be assigned the wrong stress (375 tokens).

words ending with diphthongs (incorrect %)			words ending with long vowels (incorrect %)			words with prefix + root (incorrect %)		
classify	19	76%	*absolute*	22	88%	*agriculture*	25	100%
detail	16	64%	*concrete*	25	100%	*infamous*	25	100%
fertile	18	72%	*contribute*	25	100%	*interview*	25	100%
imitate	20	80%	*obsolete*	25	68%	*triangle*	25	100%
marmalade	25	100%	*substitute*	20	80%	*vegetable*	13	52%
total	**98**	**78%**		**117**	**94%**		**113**	**90%**

4.1.3. Incorrect assignment on words with suffixation morphemes

Most of the suffixation morphology combinations illustrated here are verbs that are converted into nouns or adjectives (see Table 3). Some are nouns, such as *democrat*, *diplomat*, *photo*, and *policy* that are converted into adjectives. The word *mechanism* is a noun, from which the two adjectives *mechanic* and *mechanical* are derived.

Word groups (1) and (2) are the ones that the participants were already familiar with, and which we expected would be read correctly at a rate of 100%. However, only 80% of the participants got group (1) words right, while 93% got group (2) words right. The subjects still have a problem with the five words *alternate*, *democrat*, *diplomat*, *predicate* and *politics*. Only 3 out of 25 got *politics* right. Instead of saying/'pɒlətɪks/, they placed the stress on the second syllable and pronounced it as [pə'lɪtɪks]. It seems that the subjects' mispronunciation is reasonable, because suffixes, such as *-ic* and *-ics*, often cause the stress to shift to the syllable immediately preceding them, as in the words *economic*/ˌiːkəˈnɒmɪk/and *economics*/ˌiːkəˈnɒmɪks/. The next problematic words were *alternate* and *predicate*. Only 20% of the participants got them right; the rest mispronounced them as [ˌɔːltəˈneɪt] or [ˌpredɪˈkeɪt] respectively. The reason for placing the stress on the last syllable is that the CLs are accustomed to shifting the stress to the final syllable, either for a diphthong as in *alternate* and *predicate* or in a monophthong as in *democrat* and *diplomat*. These are trisyllabic words.

The misplaced stress in the group (1) words *alternate*, *predicate*, *democrat*, and *diplomat* happens to be the correct placement in the derivational words *alternation*, *predication*, *democratic*, and *diplomatic*. In addition, having learned the words *nation* and *education*, or the words *historic* and *energetic*, the students can easily find the correct placement of the stress in the derivational words. However, a problem arises when the regularity of the stress placement is broken, and when the learners are not in the habit of consulting a dictionary to confirm the correct stress forms. Thus, no participants achieved correct pronunciations of all the derivational words in word group (3).

No matter what kind of derivational morphology combinations these words represent, there is a tendency that an incorrect stress assignment in these derivational words will be generated on the basis of the known words. Take *comparable*, for example. Students may first learn its verb form *compare*/kəm'peə/, and later the noun form *comparison*/kəm'pærɪsən/and the adjective form *comparative*/kəm'pærətɪv/. All three of these words have the stress on the second syllable. Therefore, the students may infer that the stress in *comparable*/'kɒmpərəbl/still falls on the second syllable, so they misread it as [kəm'pærəbəl]. Other mispronounced words, such as *admirable* [əd'maɪrəbəl] (for/'ædmərəbl/), *alternative* [ɔːltə'neɪtɪv] (for/ɔːl'tɜːnətɪv/), *imaginative* [ɪˌmædʒɪ'neɪtɪv] (for/ɪ'mædʒɪnətɪv/), and *predicative* [ˌpredɪ'keɪtɪv] (for/prɪ'dɪkətɪv/), result from a similar assumption that the derivational words are pronounced in the same way as the derived nouns as they had learned earlier. The word *mechanism*/'mekənɪzm/is wrongly pronounced as [mɪ'kænɪzəm] on the assumption of *mechanic*/mɪ'kænɪk/ and *mechanical*/mɪ'kænɪkl/, because students learn these two words before they learn the word *mechanism*.

The inference to assign stress in derivational words based on previously learned words can also be seen the words *democracy* [ˌdemə'kræsɪ], *diplomacy* [ˌdɪplə'mæsɪ], and *political* ['pɒlɪtɪkəl], which are based on the words *democratic*/ ˌdemə'krætɪk/, *diplomatic*/ˌdɪplə'mætɪk/, and *policy*/'pɒlɪsi/or *politics*/'pɒlətɪks/respectively. In short, the subjects' incorrect assignment of the stress is partly on the basis of the words they are familiar with. They are not aware of the complexity of the stress rules in English, and they do not consider the suffixation that may change the stress placement in a word.

More examples are the words *photograph*, *photography*, and *photographic*. When the CLs come across the word *photography*, they presume that it has the same stress pattern as the compound *photograph*, not realizing that after suffixation with '-y'", the stress is always on the antepenultimate syllable. In addition to *-graphy*, words ending with *-ian*, *-ic*, *-ical* frequently have the stress shifted to the syllable preceding these suffixes. In these, as in many other words in English, a change of suffix not only brings about a shift in stress but also a change in the accompanying vowel reduction or neutralization in the unstressed syllables.

Table 3
Words with suffixation morphemes (950 tokens).

words (1) correctness %			**words (2) correctness %**			**words (3) incorrectness %**		
admire	25	100%	*admiration*	25	100%	*admirable*	25	100%
alternate (n.)	5	20%	*alternation*	25	100%	*alternative*	25	100%
compare	25	100%	*comparison*	25	100%	*comparable*	25	100%
			comparative	25	100%			
democrat	10	40%	*democratic*	25	100%	*democracy*	25	100%
diplomat	14	56%	*diplomatic*	25	100%	*diplomacy*	25	100%
imagine	25	100%	*imagination*	25	100%	*imaginative*	25	100%
predicate	5	20%	*predication*	25	100%	*predicative*	25	100%
prefer	25	100%				*preference*	25	100%
refer	25	100%	*reference*	25	100%			
relate	25	100%	*relation*	25	100%	*relative*	25	100%
mechanic	25	100%	*mechanical*	25	100%	*mechanism*	25	100%
photo	25	100%	*photograph*	25	100%	*photography*	25	100%
policy	25	100%	*politics*	3	13%	*political*	25	100%
total	259	80%		303	93%		300	100%

Table 4
Incorrect assignment in single-stressed compounds (500 tokens).

Noun + noun compounds	**Modifier + noun phrases**
blackboard (school room writing surface)	*black board* (board that is black)
greenhouse (glass building for plants)	*green house* (house that is painted green)
gentleman (a polite and considerate man)	*gentle man* (man who is gentle)
paper boy (newsvendor)	*paper boy* (model made of paper)
English teacher (teacher who teaches English)	*English teacher* (teacher who is English)
-ing (gerund) + noun compounds	**-ing (present participle) + noun combinations**
dining-room (room for dining)	*sliding seat* (seat that can slide)
sleeping-car (car for sleeping)	*sleeping baby* (baby who is sleeping)
running-shoes (shoes for running)	*running water* (water that is running)
smoking-room (where one can smoke)	*smoking room* (room that is on fire)
cooking butter (butter intended for cooking)	*cooking butter* (butter that is being cooked)

One interesting finding is that although the stress in the word *preference* is incorrectly assigned as [prɪ'fɜːrəns], the word *reference* is seldom mispronounced. The reason for this may be that the learners may come across the word *reference* prior to the word *refer*, and become familiar with it.

4.1.4. Incorrect assignment in compounds

The following are noun compounds and occur frequently in English (see Table 4). Normally a compound made up of two nouns takes a single stress, with the first element strongly stressed, while a phrase that looks like a noun compound but functions as a modifier plus noun sequence often has a double stress (Christophersen, 1973; Kingdon, 1958; Wells, 2006).

Single-stressed noun + noun compounds are always pronounced by the CLs as a double-stressed modify + noun phrase with a primary stress on each element. The words '*blackboard*, '*greenhouse*, and '*gentleman*, for instance, are pronounced as 'black 'board, 'green 'house, and 'gentle 'man. In addition, single-stressed *-ing + noun* compounds, such as '*dining-room*, '*sleeping-car*, '*running-shoes*, are double-stressed by the subjects as 'dining-'room, 'sleeping-'car, 'running-'shoes. In fact, the -ing form is a gerund which expresses the purpose of the noun. If '*running-shoes* is read as 'running-'shoes, the meaning will be that the shoes are running, and the result will be quite ridiculous. The stress of these words always falls on the first element, and the second element receives secondary stress. No matter whether it is a single-stressed compound or an -ing + noun compound in the target language, they are considered by the Chinese learners as two words, as in [xei55 pæn214] (*blackboard*) and [ts^{h}æn55 t^{h}iŋ55] (*dining-room*); therefore, each of the two elements is given a separate stress. The reason for this tendency can often be traced back to the fact that compounds of two full-tone syllables with a single stress do not exist in the Chinese language, while "compounds with neutral tone in Mandarin only take a proportion of as small as 14% of all compound words" (Yin, 1989).

4.2. Inappropriate treatment of unstressed syllables in words

This section examines the 25 subjects' production of unstressed syllables in isolated words. The stimuli are 8 words of high occurrence (see Table 5), and the CLs are quite familiar with them and basically have no difficulty in identifying the location of the lexical stress. The CLs' major problem lies in the improper treatment of the weak vowels in the unstressed syllables. In the NS's speech, such vowels are reduced to/ə/or/i/, and are weak in terms of intensity and perception. This is in stark contrast to the CLs, who tend to maintain full vowel quality in the unstressed syllables, and thus manifest an almost complete lack of knowledge and awareness of English vowel reduction.

The stimuli are divided into four categories. The first category contains two stimuli, *broken* and *cotton*. In the NS's speech, the weak syllables only contain the syllabic nasals, but in reality schwa is also allowed in these two unstressed syllables, and perceptually it is rather weak. In our research, although many CLs' production belongs to the second type, it is often the case that their schwa manifests full vowel quality and it sounds quite strong and salient.

Table 5
The stimuli and the corresponding NS's and CLs' pronunciation.

Category	**Stimuli**	**NS**	**CLs**
1	broken	['brəʊkŋ] or ['brəʊkᵊn]	['brəʊkən]
	cotton	['kɒtn] or ['kɒtᵊn]	['kɒtən]
2	forget	[fə'get]	[fo'get]
	complete	[kəm'pliːt]	[kʌm'pliːt]
3	little	['lɪtl]	['lɪtəʊ]
	apple	['æpl]	['æpəʊ]
4	greatly	['greɪtli]	['greɪtlei]
	provide	[prə'vaɪd]	[prəʊ'vaɪd]

Table 6
The differences between the two groups with respect to the intensity prominence of the unstressed syllables.

Word	CLs (mean)	Standard deviation	NS (test value)	t	p
broken	91.00	4.86	81.43	9.85	<0.05
cotton	96.37	5.97	89.55	5.70	<0.05

What follows is an experiment investigating the differences between the two groups of subjects regarding the intensity prominence of the unstressed syllables in *broken* and *cotton*, which refers to the proportion of the intensity peak of the unstressed syllable to that of the stressed syllable. The higher the proportion is, the more prominent the unstressed syllable sounds. If the proportion turns out to be 100% or higher, this means that the unstressed syllable is as prominent as or even more prominent than the stressed syllable. As Table 6 shows, compared with the NS, the CLs' production of the unstressed syllables is significantly more prominent.

The above analysis focuses on the aspect of intensity. Additionally, the CLs' failure to make the vowel in the unstressed syllable as weak as possible is also reflected by full vowel quality. That is to say, in the unstressed syllable produced by the CLs, what we often hear is not a schwa but another vowel, such as [o], [ʌ], or even a diphthong. This fact demonstrates that the CLs do not have awareness of English vowel reduction in unstressed syllables. The next three categories are also analyzed from this perspective.

The second category has two stimuli, *forget* and *complete*. In the NS's production, the vowel in these two unstressed syllables is the schwa [ə]. According to our research, many subjects replaced [ə] with [o] in *forget* and with [ʌ] in *complete*. As shown in Fig. 2, 12 out of 25 subjects replaced [ə] with [o] in *forget* (marked by the grey spots). For example, in CL-25's production of the weak vowel, the F1 is 501 Hz and the F2 is 1032 Hz (based on the stable period of the two formants marked in Fig. 3), suggesting that the vowel in the unstressed syllable is not schwa but [o] instead.

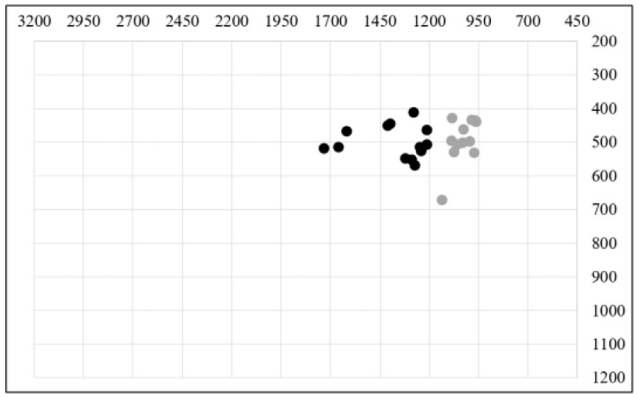


Fig. 2. The CLs' vowels in the unstressed syllables of *forget*.

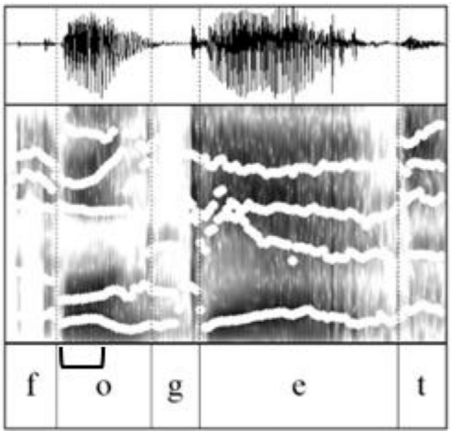


Fig. 3. CL-25's production of *forget*.

The third category has two stimuli, *little* and *apple*. In the NS's production, the unstressed syllable only has one segment, that is, the syllabic [l], based on our research; nevertheless, it was often replaced with the diphthong [əʊ] by the CLs. We calculated the F1 and F2 values of the starting (marked by the black spots) and ending (marked by the hollow spots) points of the stable period following [t]. Fig. 4 demonstrates an obvious tendency to move from the relatively central area towards the upper right corner, indicating that the tongue gradually moves from the central area towards the high and back position. For example, in CL-8's production (Fig. 5), both F1 and F2 drop gradually, with F1 moving from 874 Hz to 362 Hz and F2 from 1556 Hz to 986 Hz, indicating that the segment following [t] is not the syllabic [l] but [əʊ] instead.

The fourth category also has two stimuli, *greatly* and *provide*. In the NS's production, the vowels in the two unstressed syllables are [i] and [ə] respectively. However, the CLs have a propensity to use [ei]/[eɪ] in the word final position and [əʊ] in the word initial position. Following the above approach, we have also calculated the F1 and F2 values of the starting (marked by the black spots) and ending (marked by the hollow spots) points of the stable period of the vowels in the unstressed syllables. Fig. 6 shows an obvious tendency to move from the relatively central area towards the upper left corner, indicating that the tongue gradually moves from the central area towards the high and front position. For example, in CL-21's production (Fig. 7), the F1 and F2 speckled lines gradually repel each other, with F1 moving from 462 Hz to 249 Hz and F2 from 2231 Hz to 2664 Hz, suggesting that the vowel is not [i] but [ei] instead.

4.3. Inappropriate prominence to syllables in sentences

At the sentence level we test the subjects' performance by asking them to read 3 sentences containing both rising and falling nuclear tones and with a variety of placement of nuclear syllables. Also, the sentences contain the stressed syllable, the unstressed syllable, and the weakly stressed syllable of the function words. Three sentences are given with intonation phrase boundary marks:

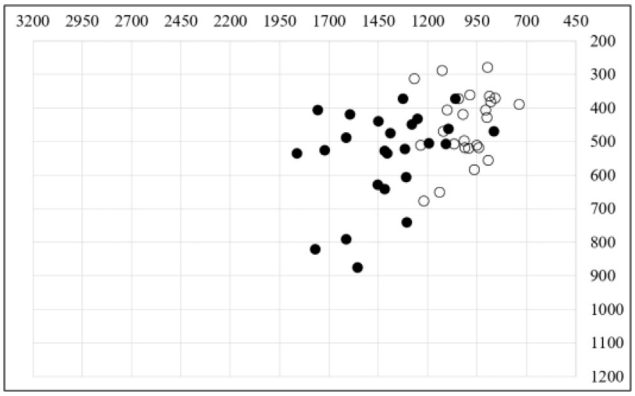


Fig. 4. The CLs' vowels in the unstressed syllables of *little*.

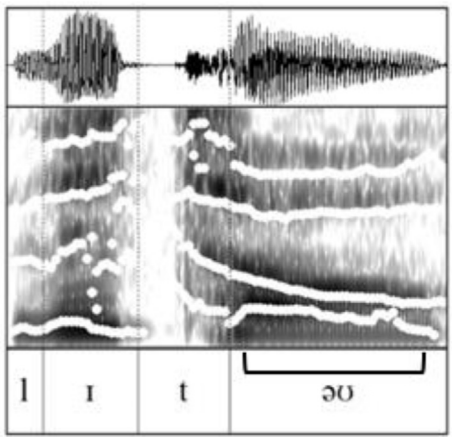


Fig. 5. CL-8's production of *little*.

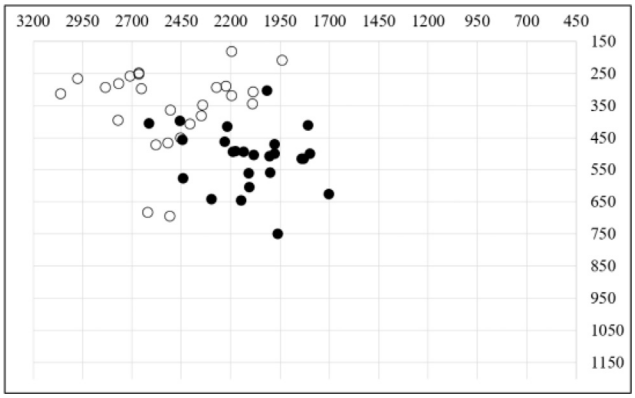


Fig. 6. The CLs' vowels in the unstressed syllables of *greatly*.

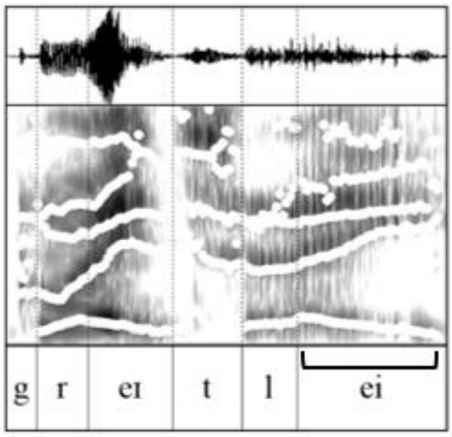


Fig. 7. CL-21's production of *greatly*.

(1) *We could go by bus.* ||
(2) *Please report to the enquiry desk.* ||
(3) *You can pick it,* | *you can seal it,* | *and you can load it.* ||

The first as well as the second sentence is co-extensive with one intonation phrase (IP). The last sentence is divided into three IPs. According to the conventional regularities of the sentence stress, the three sentences are supposed to be read as:

(1) *We could 'go by* ***bus**. ||
(2) *'Please re'port to the en****quir**y ˌ*desk.* ||
(3) *You can*/***pick** it,* | *you can* /***seal** it,* | *and you can* ***load** it.*

The bold and underlined part is the nuclear syllable. They are indicated by falling or rising tones, and should be the most prominent syllable in each IP. In other words, sentence (1) has its nuclear syllable, i.e., the most prominent stressed syllable, on the last content word, which is the normal placement for nuclear stress. The deaccenting of the 'last content word' in sentence (2) is the weaker-stressed part of a compound.[3] Function words in sentence (3) are not to be stressed, and thus, the three nuclei should be placed on *pick*, *seal*, and *load*. In addition, the nuclear tones for sentence (3) are rise + rise + fall, a typical pattern of a listing tone for a sequence of items.

[3] Most compounds in English are single-stressed, that is, the main lexical stress goes on the first element.

Table 7
The differences between the two groups with respect to the pitch prominence of each syllable.

Syllable	CLs (mean)	NS
we	42.58	-10.94
could	15.90	-4.17
go	**4.99**	**17.19**
by	4.96	6.25
bus	**13.17**	**50.00**
please	**33.72**	**44.83**
re-	-2.45	24.54
-port	**21.42**	**27.86**
to	3.42	27.12
the	-8.21	17.90
en-	-9.46	17.90
-quir-	**28.13**	**50.00**
-y	-7.96	-17.90
desk	-4.01	-44.10
you	9.07	-20.18
can	6.51	-13.74
pick	**37.07**	**-34.80**
it	7.88	50.00
you	1.34	-16.89
can	7.81	-12.91
seal	**41.90**	**-28.15**
it	-7.84	50.00
and	25.19	20.37
you	16.52	17.78
can	32.73	32.22
load	**34.12**	**50.00**
it	-15.88	-37.04

4.3.1. Pitch prominence

The experiment tests the differences between the two groups of subjects with respect to the pitch prominence of each syllable. To be more specific, the pitch prominence of a certain syllable (in percentage %) refers to the proportion of the difference between the maximum pitch value of that syllable and the median pitch value of the corresponding IP, to the whole pitch range of that IP. On the one hand, the higher the proportion is, the more prominent the syllable is; the proportion of the most prominent syllable in a certain IP should reach as high as 50%. On the other hand, a negative value means that the syllable is not prominent, resulting from the fact that even the maximum pitch value of that syllable turns out to be lower than the median value. The data are shown in Table 7. Both stressed and nuclear syllables are in bold. The nuclear syllables are underlined.

With regard to the pitch prominence of the stressed (*go, please, -port*) and nuclear syllables (*bus, -quir-, pick, seal, load*), most CLs made these syllables prominent by using higher pitch. The CLs are significantly different from the NS in terms of two perspectives. First, in the unmarked sense, the pitch of the nuclear syllables of *pick* and *seal* initiates nearly from the bottom of the pitch range, and then rises a bit. Therefore, in the NS's speech even the maximum values of those two syllables are much lower than the median values of the corresponding two IPs, with the prominence of *pick* reaching as low as −34.80% and *seal* −28.15%. In contrast, the CLs show an obvious tendency to use falling pitch movement to mark these two nuclear syllables as indicated by the positive values of pitch prominence. Second, in spite of the fact that the CLs often used higher pitch to mark the rest of the stressed and nuclear syllables as the NS has done, those produced by the CLs appear to be less prominent if compared with the NS's.

Next, the CLs' major problem lies in the pitch prominence of the prehead, the part beginning from the first unstressed syllable to the last unstressed one just preceding the first stressed syllable of a certain IP. The NS made those syllables as weak as possible, which is suggested by the negative values of prominence. For example, even the maximum pitch values of *we* and *could* are lower than the median value of the IP, and in consequence they are much less prominent (−10.94% and −4.17%) than the stressed counterparts of the same IP (see Fig. 8). In the same vein, as for the first two IPs of the last sentence, the syllables *you* and *can* are much less prominent, with each being lower than −10%. Nevertheless, these unstressed syllables are not properly produced by the CLs, whose production appears to be quite prominent. For the syllables *we* and *could*, the CLs' prominence reaches as high as 42.58% and 15.90% respectively. CL-10's production as shown in Fig. 9 represents a typical example of this kind. The CLs' preheads of the first two IPs of the last sentence are still prominent. Although their prominence is generally lower than 10%, they are much more prominent if compared with the NS's prominence. The NS's prehead of the last IP of the same sentence is relatively prominent, and it gradually moves from the mid-high to high level, indicating a certain degree of emphasis. As a result, the CLs are similar to the NS in this regard.

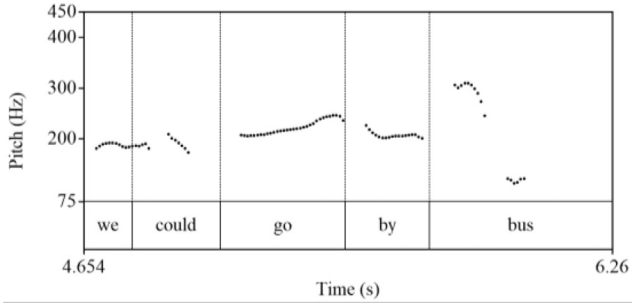


Fig. 8. NS's production of sentence (1).

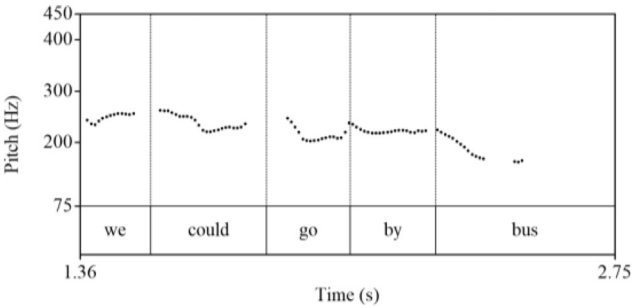


Fig. 9. CL-10's production of sentence (1).

The second problem lies in the pitch prominence of the head, the part beginning from the first stressed syllable to the syllable just preceding the nucleus. The pitch of the unstressed syllables in the head is in large measure reliant upon the preceding stressed ones. To be more precise, the unstressed syllables should be maintained at almost the same pitch level as the preceding stressed syllable, and the whole pitch movement of the head manifests a gradual downward trend. This pattern is shown in the NS's speech. Nevertheless, in the production of many CLs, the pitch movement of the head is not well controlled, and consequently, it shows an inappropriate up-and-down fluctuation pattern.

The unstressed syllables in the tail, the part following the nuclear syllable, do not seem to be that problematic for the CLs with respect to pitch prominence, because the compressed pitch range of the tail is a cross-linguistic phenomenon, and there is substantial evidence that it exists in Chinese and English. Therefore, in both the NS's and the CLs' production, the tail following the falling pitch movement of a certain nuclear syllable tends to be seriously compressed.

4.3.2. Duration prominence

The experiment also tests the differences between the two groups of subjects in terms of the duration prominence of each syllable. Specifically, the duration prominence of a certain syllable (in percentage %) refers to the difference between the proportion of that specific syllable to the IP and the average proportion of each syllable of that IP. On the one hand, the more obvious the difference is, the more prominent that syllable is. On the other hand, if the value turns out to be negative, it means that the syllable is not prominent enough. The data are shown in Table 8.

Regarding the duration prominence of the stressed (*go, please, -port*) and nuclear syllables (*bus, -quir-, pick, seal, load*), most CLs made them more prominent, but the prominence appears to be less satisfactory if compared with that of the NS. For example, in the first sentence many CLs made *go* less prominent; as a result, its duration prominence value turns out to be negative, being just −0.17%. However, the NS's prominence reaches as high as 3.63% (see Fig. 10). CL-14's production as shown in Fig. 11 represents a typical example of this kind. For the syllables *bus*, *please*, *-quir-*, *pick*, and *seal*, the CLs' production is significantly less prominent than that of the NS.

In terms of the duration prominence of the unstressed syllables, the CLs' performance is generally acceptable. They managed to make them quite short like the NS, despite the fact that in some cases significant differences still exist between the two groups of subjects. For instance, the statistical analysis shows that the NS and the CLs are significantly different from each other in the duration prominence of *we, could, to* and etc.

5. Conclusion

The study has found that CLs produced unstressed syllables and weakly-stressed function words with a significantly higher pitch, a longer duration, and greater intensity than a native English speaker. Consequently, they showed less

Table 8
The differences between the two groups with respect to the duration prominence of each syllable.

Syllable	CLs (mean)	NS
we	-5.72	-10.86
could	-1.53	-3.33
go	**-0.17**	**3.63**
by	-4.08	-4.14
bus	**11.50**	**14.70**
please	**4.74**	**7.40**
re-	-4.41	-4.42
-port	**2.33**	**3.10**
to	-2.72	-6.80
the	-4.88	-4.92
en-	-3.30	-4.34
-quir-	**4.85**	**7.51**
-y	-4.97	-6.84
desk	8.36	9.32
you	-10.19	-14.60
can	-1.84	-1.62
pick	**3.78**	**6.30**
it	8.25	9.92
you	-12.16	-15.43
can	-5.53	-4.54
seal	**13.10**	**19.91**
it	4.59	0.07
and	-6.04	-10.17
you	-5.66	-3.79
can	-2.15	2.73
load	**6.74**	**7.57**
it	7.10	3.66

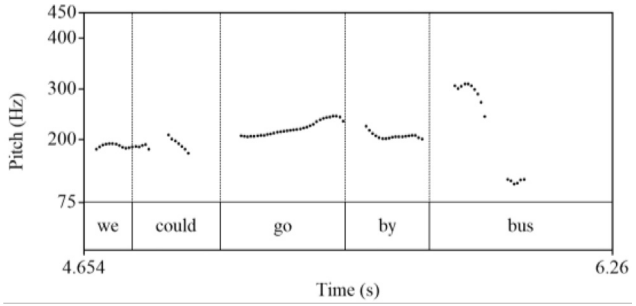


Fig. 10. NS's production of sentence (1).

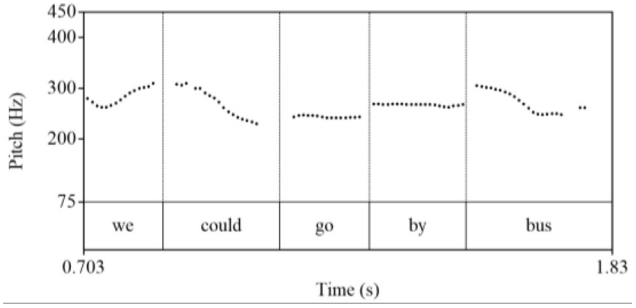


Fig. 11. CL-14's production of sentence (1).

differentiation between stressed and unstressed syllables than the native speaker in terms of syllable duration, pitch prominence, and vowel reduction. The differences between the two languages affect how the Chinese EFL learners assign stress and pitch prominence.

With respect to the assignment of lexical stress, suffixes were more likely to be pronounced incorrectly by the CLs because they are deeply affected by the words they have learnt and tend to put the stress on the original stress position, without taking the stress shifts caused by derivation into consideration. Single-stressed noun + noun compounds are always pronounced by the CLs as double-stressed modify + noun phrases with a primary stress on each element. No matter whether the word is a single-stressed compound or an −ing + noun compound in the target language, they are considered by the CLs to be two words, and therefore each of the two elements is given a separate stress. The reason for this tendency can be traced to the fact that real compounds with a single stress do not exist in the learners' language. In fact, the CLs' attempts to formulate rules for stress may be completely governed by the consideration of simplicity and unmarkedness; in practice, they may ignore much of the target language data and devise a stress system that is as simple and unmarked as possible.

Regarding the isolated words, the CLs have a tendency to give inappropriate prominence to the unstressed syllables. Specifically, their major problem lies in the improper treatment of weak vowels in those syllables. In the NS's speech, such vowels are reduced to/ə/or/i/and are weak in terms of their intensity and perception. This is in sharp contrast with the CLs, who tend to maintain a full vowel quality in the unstressed syllables, and thus show an almost complete lack of knowledge and awareness of English vowel reduction.

As far as the pitch prominence of sentences is concerned, the CLs' control of pitch of unstressed syllables in preheads and heads is not satisfactory. The preheads do not start at a low pitch and the pitch is often maintained at a high level, making their speech sound unnatural. In addition, the pitch of unstressed syllables in heads is not properly controlled with reference to that of the preceding stressed syllable, but shows an inappropriate up-and-down fluctuation pattern. Regarding stressed and nuclear syllables, generally speaking the Chinese EFL learners failed to make them prominent enough.

In terms of the duration prominence, to some extent the CLs' performance is acceptable. However, it is less satisfactory regarding stressed and nuclear syllables. Most of the CLs' production of the vowel in the weak form of the function words is less reduced than that of the NS. Thus, the durations of the stressed syllable, the unstressed syllable, the nuclear syllable, and the weakly stressed syllable of the function words are not significantly different in the CLs, making their speech less diverse, less complex, and thus more like Chinese.

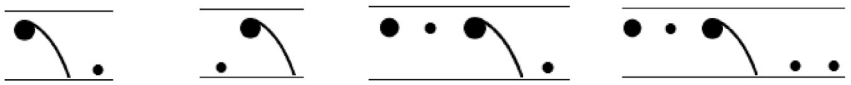
6. Pedagogical implications

This and other experimental data from L2 stress acquisition studies show that native Chinese speakers have no phonological knowledge of hearing or perceiving word stress (Altman, 2006; Wayland & Guion, 2004), because the intrinsic phonology of the Chinese language enables speakers to judge the tone, but does not allow them to perceive stress (Wang & Feng, 2006). To overcome the phenomenon of "stress-deafness" (Peperkamp & Dupoux, 2002), Chinese EFL learners need a systemic training in English stress.

The placement of word stress in English can occur on any syllable, depending in part on the origin of the word, affixes, and the word's grammatical function in an utterance. This apparent lack of predictability as to where the stress falls is confusing to CLs. Therefore, explicit teaching and learning of word stress patterns should be a part of the pronunciation curriculum. Misplacement of stressed syllables is common, particularly the addition of derivational affixes to a word. Such mistakes can be avoided only by acquiring sound knowledge about word stress patterns in English, and by remembering the different lexical stress patterns of polysyllabic words such as '*photograph*, *pho*'*tographer*, and ˌ*photo*'*graphic*.

Incorrect placement of a compound stress not only hinders intelligibility, but is also accompanied by an inappropriate pitch pattern, which is a very significant source of error, even for advanced learners. Therefore, knowing a few guidelines can make compound stress easier to learn. The best thing to do is to practice, and pay particular attention to the stress in all new compounds until learners can produce them subconsciously, thus getting a "feel" for the stress patterns in English compounds.

The difference between stressed and unstressed syllables is more significant in English, while the Chinese language is spoken with syllable timing, i.e. one syllable, one word, and one tone. Thus, each syllable has nearly the same degree of loudness and duration. Therefore, to achieve correct pronunciation of a word stress, Chinese learners of English need to know how native speakers highlight a stressed syllable with loudness, length, and pitch variation, and how they produce unstressed syllables with vowel reduction. Teachers should use exercises which can develop students' awareness of weak syllables, as well as exercises on prominent word stress and sentence stress. Learners should be trained to practise saying the pitch patterns of English words with alternative use of primary, secondary, and unstressed syllables. For example.



After they have received enough training on pitch patterns at word level, learners need to move on to pitch patterns at phrasal and sentential level until they can use these rhythmic patterns automatically.

The negative impact of Chinese EFL learners' unnatural sentence stress in English comes from habits which are formed through years of learning English. In this aspect, fossilization does exist in the Chinese EFL learners' production of English sentence stress. Thus, pedagogical implications should include distinguishing content words from function words, being aware of the modification of major sentence stress in certain contexts, giving prominence only to the words that are important, and weakly stressing the words that are unimportant in a sentence. Moreover, they need to develop the identification and production of weak forms. The best way to learn English sentence stress and weak forms is to put utterances in situational contexts.

As sentence stress is closely linked with intonation, it frequently undergoes certain modifications signaled an increase of duration, pitch, and/or prominence. In connected speech, the same word may be stressed or unstressed depending upon the context of an utterance and the degree of emphasis or contrast placed on a certain part of a sentence. The more important a word is, the stronger the stress it receives. Therefore, it is necessary to distinguish some levels of stress used in sentences, such as major sentence stress, emphatic stress, contrastive stress, and new information stress.

Acknowledgements

This is one of the phased achievements of a larger project (*The Study of the English Pronunciation Problems of Chinese EFL Learners*) funded by *The National Social Science Fund of China* (17BYY206), whose support is gratefully acknowledged. The research is chaired and conducted by Youhong Bu.

References

Altman, H. (2006). *The perception and production of second language stress: A cross-linguistic experimental study*. PhD dissertation. University of Delaware.
Beckman, M. E. (1986). *Stress and non-stress accent. (Netherlands phonetic archives No. 7)*. Second printing, 1992. In Walter de Gruyter (Ed.). Dordrecht: Foris.
Berinstein, A. (1979). *A cross-linguistic study on the perception and production of stress. UCLA working papers of phonetics* (vol. 47). Los Angeles: UCLA Phonetics Laboratory.
Bi, R., & Chen, H. (2013). A Longitudinal study of Chinese English language learners' tone patterns. *Journal of Foreign Languages and their Teaching, 1*, 50–54.
Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer* Version 6.0.55. http://www.fon.hum.uva.nl/praat/. (Accessed 15 June 2019).
Bolinger, D. L. (1958). A theory of pitch accent in English. *Word, 14*, 109–149.
Campbell, N., & Beckman, M. (1997). Stress, prominence, and spectral tilt. In A. Botinis, G. Kouroupetroglou, & G. Carayiannis (Eds.), *Proceedings of ESCA workshop on intonation: Theory, models and applications* (pp. 67–70) (Athens).
Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge: Cambridge University Press.
Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge: Cambridge University Press.
Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California.
Chen, H. (2008). *On Chinese EFL learner's English intonation patterns*. Shanghai: Shanghai Foreign Language Education Press.
Chen, Y., Robb, M. P., Gilbert, H. R., & Lerman, J. W. (2001). A study of sentence stress production in Mandarin speakers of American English. *Journal of the Acoustical Society of America, 109*, 1681–1690.
Christophersen, P. (1973). *An English phonetics course*. London: Longmans.
Clark, J., & Yallop, C. (1995). *An introduction to phonetics and phonology* (2nd ed.). Oxford: Blackwell.
Cooper, W., & Sorensen, J. (1981). *Fundamental frequency in sentence production*. New York: Springer-Verlag.
Cruttenden, A. (2008). *Gimson's pronunciation of English* (7th ed.). London: Hodder Education.
Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge, UK: Cambridge University Press.
Cutler, A. (2015). Lexical stress in English pronunciation. In M. Reid, & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 106–124). Malden, MA: Wiley Blackwell.
Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *Tesol Quarterly, 39*(3), 399–423.
Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America, 27*, 765–768.
Fry, D. B. (1958). Experiments in the perception of stress. *Language and Speech, 1*, 126–152.
Fry, D. B. (1965). The dependence of stress judgments on vowel formant structure. In E. Zwirner, & W. Bethge (Eds.), *Proceedings of the 5th international congress of phonetics sciences* (pp. 306–311) (Basel).
Gao, W., Xu, Y., & Mu, F. Y. (2015). An experimental study on rhyme-focused teaching for Chinese EFL learners. *Journal of Foreign Language Teaching and Research, 6*, 861–873, 960.
Guo, X., & Chen, X. (2017). On production of English lexical stress by Beijing dialect speakers and Cantonese speakers, 02 *Foreign Language Teaching and Research, 49*, 188–201, 319.
Heselwood, B. (2005). *Phonetics and phonology of English: Lecture handouts. Department of linguistics and phonetics*. Leeds, UK: University of Leeds.
Ho, A. (1976). Mandarin tones in relation to sentence intonation and grammatical structure. *Journal of Chinese Linguistics, 4*, 1–13.
Johnson, K. (2012). *Acoustic and auditory phonetics* (3rd ed.). UK: Wiley-Blackwell.
Kijak, A. (2009). *How stressful is L2 stress? A cross-linguistic study of L2 perception and production of metrical system*. PhD dissertation. Untrecht: LOT.
Kingdon, R. (1958). *The groundwork of English stress*. London: Longmans.
Ladefoged, P. (2003). *Phonetic data analysis*. London: Blackwell Publishing Ltd.
Ladefoged, P., & Johnson, K. (2015). *A course in phonetics* (7th ed.). USA: Cengage Learning.
Lewis, C. (2017). *Word stress in non-native English: Evidence from English as a lingua franca*. http://fass.ubd.edu.bn/news/lewis.html. (Accessed 5 July 2019).
Lewis, C., & Deterding, D. (2018). Word stress and pronunciation teaching in English as a Lingua Franca contexts. *CATESOL Journal, 30*(1), 161–176.
Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America, 32*, 451–454.
Lieberman, P. (1975). *Intonation, perception, and language*. Cambridge: MIT Press.
Meng, X. J., & Wang, H. M. (2009). Boundary tone patterns in Chinese English learners' read speech. *Journal of Foreign Language Teaching and Research, 6*, 447–451, 960.
O'Connor, J. D., & Arnold, G. F. (1973). *Intonation of colloquial English* (2nd ed.). London: Longman.
Peperkamp, S., & Dupoux. (2002). A typological study of stress 'deafness'. In C. Gussenhoven, & N. Warner (Eds.), *Laboratory phonology 7* (pp. 203–240). Berlin: Mouton de Gruyter.
Pike, K. (1948). *Tone languages*. Ann Arbor: University of Michigan Press.
Potisuk, S., Gandour, J., & Harper, M. (1996). Acoustic correlates of stress in Thai. *Phonetica, 53*, 200–220.
Roach, P. (2009). *English phonetics and phonology: A practical course* (4th ed.). Cambridge, UK: Cambridge University Press.

Tench, P. (1996). *The intonation systems of English*. London: Cassell.
Wang, G. Z. (2011). *English pronunciation and intonation practice* (2nd ed.). Beijing: Higher Education Press.
Wang, Z. J., & Feng, S. L. (2006). Tonal contrast and disyllabic stress patterns in Beijing Mandarin. *Linguistic Sciences, 5*(1), 3–22.
Wayland, R. P., & Guion, S. G. (2004). Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language and Learning, 54*(4), 681–712.
Wells, J. C. (2006). *English intonation: An introduction*. Cambridge, UK: Cambridge University Press.
Wells, J. C. (2008). *Longman pronunciation dictionary* (3rd ed.). Harlow, UK: Pearson Longman.
Xu, X. M. (2013). Tone-based rhythm of Chinese and stress-based rhythm of English. *Journal of PLA University of Foreign Languages, 5*, 1–5.
Xu, Y. M. (2019). *Contrastive studies of sound systems in English and Chinese*. Beijing: Foreign language Teaching and Research Press.
Xu, Y. M., & Shen, J. X. (2016). Phonological differences between English and Chinese stress. *Journal of Foreign Language Teaching and Research, 5*, 643–656, 799.
Yang, J. (2012). A Study on the English stress placement of Chinese EFL learners. In *The third international symposium on tonal aspects of languages*. http://www.isca-speech.org/archive. (Accessed 5 July 2019).
Yin, Y. (1989). *Phonological aspects of word formation in Mandarin Chinese*. PhD dissertation. University of Texas.
Zhang, Y. H. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *Journal of the Acoustical Society of America, 123*, 4498–4513.

Youhong Bu is Professor of Phonetics at Foreign Languages College, Shanghai Normal University, Shanghai, China. She is an expert of *Shanghai Education Evaluation Institution*, an executive member of *China's Pronunciation Teaching & Phonetics Research Committee*, and Chief English expert of *National Teacher Training Program* at Shanghai Normal University. Her current interests lie in English phonetics and phonology, L2 pronunciation learning and teaching. She has published extensively in these fields in such core academic journals as *Journal of Foreign Language Teaching and Research*, *Journal of Technology Enhanced Foreign Language Education*, and *Journal of E-education Research*. She has published 7 works and 85 tapes on English pronunciation and intonation, listening comprehension, junior and senior middle school English teaching in *Beijing Foreign Languages Audiovisual Publishing House* and *Gansu Audiovisual Textbook Publishing House*. She has been granted several awards, the first prize winner of *Shanghai Teaching Achievements Awards* (2017), and the second prize winner of *National Teaching Achievements of MA Degree* (2018). In 2016, she won the award of the *Excellent Teaching Management Worker in National MA Education*. In 2012, she won *Shanghai Excellent Teacher Awards*. She is recognized as one of China's outstanding representatives of “Received Pronunciation” in English phonetics circle at home and abroad.

Zhengdong Zhou is a PhD student at College of Foreign Languages and Literature, Fudan University, Shanghai, China. His current interests include the phonological study of Chinese dialects and the acoustic study of Chinese learners' English pronunciation.