Helping Cantonese ESL learners overcome their difficulties in the production and perception of English speech sounds

Alice Y.W. Chan
Department of English
City University of Hong Kong
83 Tat Chee Avenue
Kowloon
Hong Kong

Email: enalice@cityu.edu.hk
Fax: 852-3442-0288
Tel: 852-3442-9752

Keywords: speech production and perception second language learning pronunciation difficulties pronunciation teaching program
Abstract

Cantonese ESL learners’ speech production and speech perception abilities have great impact on their learning of English pronunciation. This article attempts to give a brief summary of documented speech production and perception difficulties encountered by Cantonese ESL learners in Hong Kong and provide some principles which underlie effective pronunciation teaching and learning. It is suggested that observable articulatory features should be highlighted, subtle differences between confusable sound pairs should be focused on, and mother-tongue interference should be tackled. It is also argued that perceptual training is one inevitable component of an L2 pronunciation program going hand in hand with production training.
Introduction

Much previous literature on Cantonese ESL pronunciation learning focuses on learners’ production of English speech sounds (e.g. Bolton & Kwok, 1990; Chan & Li, 2000; Chan, 2005; Hung, 2000; Kenworthy, 1987; Lo, 2005; Stibbard, 2004). Learners’ perception abilities and related difficulties have received relatively less attention until very recently when this area was brought to life by researchers such as Chan (2001) and Chan (2006c, 2006d, 2007a, 2007b). Speech perception often bears an intimate relation to speech production, and it has often been claimed that learners’ perception of L2 phonetic segments may affect the accuracy with which these segments can be produced (e.g. Flege, 1995; Munro & Derwing, 1995; Schmid & Yeni-Komshian, 1999). In this article, some documented L2 production and perception difficulties encountered by Cantonese ESL learners will be summarized and suggestions on how ESL teachers can help students overcome their students’ L2 production and perception difficulties will be given.

Difficulties in the production and perception of English speech sounds encountered by Cantonese ESL learners

Production problems

In analyzing Cantonese ESL learners’ pronunciation problems, researchers have often used a standard variety of English, such as Received Pronunciation (RP), or to a lesser extent American English, as the model for analysis. Productions which conform to the norms of the standard variety adopted are regarded as accurate whereas those which deviate from the norms are often given the label “problems.” A number of production problems encountered by Cantonese ESL learners have been documented. Some of these are based on the researchers’ observations corroborated by a substantial body of earlier findings (e.g. Chan & Li, 2000), whereas many others are the empirical findings of rigorous research studies (e.g. Chan, 2006a, 2006b; Hung, 2000; Stibbard, 2004). Among the difficulties documented are devoicing of final obstruents (e.g. devoicing of [z] in words like *rose* /raʊz/,) deletion of consonants in consonant clusters (e.g. deletion of [r] in words like *Fred* /frɛd/), substitution of an L1 sound for an L2 sound, (e.g. substitution of [ts] for /tʃ/ in words like *cheap* /tʃi:p/), substitution of a non-target L2 sound for a
target L2 sound (e.g. substitution of [w] for /r/ in words like rose /roʊz/, or [e] for /æ/ in words like bad /bæd/), non-distinction of the length difference between L2 short and long vowel pairs (e.g. /u:/ and /u/ in words such as food /fuːd/ and foot /fʊt/), vocalization or omission of final dark [I] (e.g. the use of the vowel [u] to substitute for dark [I] in words like dill /dɪl/), and the like. It has often been argued in these studies that most of the problems can be attributed to mother-tongue influence, in that segments non-existent in the learners’ mother tongue, Cantonese, are often found to have caused production difficulties, whereas segments shared by both the native language and target language phonemic inventories do not pose great production difficulties. However, other factors may also play a role, such as the universal difficulty of the English dental fricatives /θ/ and /ð/ (the sounds are very rare in the world’s languages and are thus extra difficult) (Maddieson, 1984), learners’ avoidance behaviour (e.g. employing deletion in the pronunciation of consonant clusters rather than insertion in an attempt to avoid re-syllabification) and their overgeneralization of phonetic properties to unsuitable contexts (e.g. aspirating plosives after /s/ in words like star /stɑː/) (Chan, 2006b).

Perception problems

Speech production is not the only arena where Cantonese ESL learners experience problems: perceptual problems are also rampant. In speech perception research, a standard variety has also often been adopted in the preparation of research materials (i.e. stimuli for perception). In her study of the perception and production problems encountered by Cantonese ESL learners in Hong Kong, Chan (2001) found a positive correlation between perception problems and production problems. The learners who consistently confused the target consonant pairs (/v, w/, /θ, f/, /ð, d/, /z, s/ and /r, w/) in production also demonstrated perceptual confusion for the same contrast pairs, and all the perceptual errors for the target items /v, θ, δ, z, r/ were the ones corresponding to their mispronounced versions /w, f, d, s, w/ respectively. Chan (2006c, 2006d, 2007a, 2007b), however, found that though learners encountered difficulties in their perception of English speech sounds, the level of difficulty was not on a par with the level of difficulty
encountered by Cantonese ESL learners in their production of English speech sounds. Those English speech sounds documented widely in the literature as causing production difficulties (e.g. /ʃ/, /tʃ/, /æ/ and /r/) (Chan & Li, 2000; Stibbard, 2004) were not particularly difficult to perceive, whereas sounds which have not been documented as difficult for production (e.g. /e/, /f/, /l/) did create some perception problems for Cantonese ESL learners. Regarding the actual perception problems, it was also found that learners were better at discriminating isolated sounds from each other (e.g. /i:/ and /u/ in isolation) than identifying words containing contrasting sounds (e.g. minimal pairs such as bean /bɪn/ and bin /bɪn/). She argues that the perception difficulties ESL learners encounter may be due to their misconception of word pronunciation rather than their inability to discriminate acoustic differences – Because learners’ mental representation for perception may be mediated by predetermined word pronunciations and input acoustic signals may be converted to forms which fit their distorted mental representation, incorrect perceptual judgments may result. Mother tongue interference, which has been maintained as a main contributor to production problems, is argued to have played a minimal role in perception (Chan, 2007a).

What can teachers do?

The speech production and perception problems that Cantonese ESL learners encounter are ubiquitous. ESL teaching professionals are recommended to shape teaching activities to target the resolution of these obstacles to successful second language mastery. In view of the nature of the problems as well as the possible causes, I would like to put forward some principles and suggest some activities where appropriate, which may help guide the design of pronunciation teaching programs.

Unlocking the source of problems

As can be seen from the literature review, Cantonese ESL learners encounter problems both in speech production and in speech perception. The first step to rewarding pronunciation teaching is thus the unraveling of the source of learner problems: Are the problems perception-based or production-oriented? If the problems are perceptual, are they to do with discrimination or identification? Discrimination, the ability to marshal
perceptual resources to differentiate between signals in the auditory and perceptual periphery, is largely related to sensory detection; whereas identification, the ability to associate labels to a signal based on its auditory content, may be associated with pre-asserted word knowledge. Even if contrasts (e.g. voicing contrasts) get discriminated, the corresponding segments (/z/ as opposed to /s/) and words (loose /luːs/ as opposed to lose /luːz/) may not get identified unless learners are able to associate correct labels to perceived speech signals. Production problems, on the other hand, can be differentiated in terms of the ability to produce a certain phonetic feature and the ability to actualize the phonetic feature in continuous speech. Learners who are capable of producing a phonetic feature in isolation may still have difficulty in actualizing the feature in continuous speech. Therefore, the source and nature of learner problems should be carefully diagnosed and accordingly tackled so as to ensure ESL learners can gain full control over their pronunciation learning. The use of both perception diagnostic tests (e.g. teachers producing the target sounds in isolation or in minimal pairs and asking students to identify the spoken sounds or words) and production diagnostic tests (e.g. teachers showing students minimal word pairs in spelling form or in pictures and requiring students to read out the words) is necessary before the implementation of any remedial training.

*Alerting conscious attention to correct model and differentiating between confusable L2 sounds*

It has often been claimed that an important key to successful second language learning is continuous exposure to the target language. Continuous exposure to the target language is, without doubt, necessary, without which language learning is bound to fail. However, it is more important, even vital, to alert learners to the value of conscious attention paid to the correct model and deliberate differentiation between confusable L2 sounds. Many ESL learners are unaware of their own mispronunciations, the differences between their productions and the correct model, and the subtle differences between certain sound pairs, such as /iː/ and /ɪ/. These mispronunciations are simply unattended to, thus leading to fossilization. Conscious attention paid to the correct model, coupled with the deliberate differentiation between the correct model and anomalies, is facilitative of
speech perception, especially identification (see Section Unlocking the source of problems), as learners can learn to associate a correct model to a correct label. Awareness of the articulatory and acoustic differences between confusable sound pairs, together with conscious attention to the correct model, can help alleviate production problems resultant from misconception of word pronunciations. Learners who, for instance, do not readily recognize the difference between long and short vowel pairs such as /i:/ and /ɪ/ in words like cheap /tʃɪp/ and chip /tʃɪp/ can have their perception problems solved from consciously attending to the length differences between the two. It is only through rectifying their inaccurate preconception of the pronunciations of these words that learners would learn to produce them correctly. Those who are unaware of their devoicing habits in pronouncing words containing voiced obstruents, such as zip /zip/ and food /fuːd/, will attend to the voicing contrast only after having been consciously alerted to the difference between their pronunciations and the correct model. Although some researchers argue that it is more meaningful to teach pronunciation in meaningful units than to teach minimal pairs or isolated segments (e.g. Chela-Flores, 2001), a number of popular pronunciation resources used by many ESL teachers worldwide, for example, Baker (1981), also include teaching and learning materials which differentiate between confusable L2 sounds in minimal pairs.

Focusing on observable articulatory features

It is of paramount importance that ESL learners be alerted to the salient articulatory features of certain L2 sounds, especially the differences between “similar” pairs. While the articulatory features of some L2 sounds may not be easily observable, such as the raising of the back of the tongue in the production of dark [H] (Ladefoged, 2006) or the voicing of final consonants, others have clearly observable articulatory features an awareness of which may facilitate acquisition, such as the difference in the tongue height between /æ/ and /e/ (the former involving a greater degree of lowering of the jaw than the latter), the difference in the lip shape between /ʃ/ and /s/ (the former having a protruded lip shape), and the noticeable attribute in the production of labiodental sounds such as /v/ (with the teeth “biting” the lower lip). If learners are aware of
these observable articulatory features, perceptually tuning in to the presence or absence of the features in their face-to-face encounters with English speakers, discrimination and identification problems associated with confusable sound pairs may be eliminated. An awareness of salient articulatory features is also facilitative of speech production, prompting learners to consciously attend to accurate articulations. Such a consciousness-raising approach has great potential for helping learners notice their L2 production and perception errors and progressively approximate target language norms. Similar techniques have been adopted by pronunciation clinics to help ESL/EFL learners overcome their pronunciation problems and have been found to be conducive to pronunciation learning. An example is the pronunciation clinic for Asian learners of English in Australia (Wajnryb, Coan & McCabe, 1997), which used exercises aimed at raising students’ awareness of the physicality of sound production using small hand mirrors.

**Tackling mother tongue interference**

Mother tongue interference, which has been argued as one significant contributor to production difficulties, should be explicitly tackled. Substitution of Cantonese /ts/ and /dz/ for English /tʃ/ and /dʒ/ respectively without the required lip-rounding, for example, may be readily overcome by an explicit comparison between the target and substitution sounds, with particular focus on the difference between the two and, if possible, the observable articulatory features of the target L2 sounds (see Section Focusing on observable articulatory features). Learners who are used to substituting Cantonese [œ] for English /ɜ:/ with undesirable lip-rounding in words such as *bird* /bɔːd/ and *fur* /fɔː:/ will also benefit from such comparisons. It should come as a welcome enlightenment to learners if they are alerted to the differences between the mother tongue and the target language, especially if the differences are subtle but readily amenable to correction without tremendous articulatory efforts. Hung (1993) also argues for the facilitative effects of raising learners’ consciousness of the differences between a learner’s L1 and L2.

*Integrating speech perception with speech production*
As production and perception are intimately related, perceptual training should be included as one major component of pronunciation teaching. Although correct production is probably the ultimate target of the majority of pronunciation learning programs, learners’ perception needs should not be belittled, especially when the perception deficiencies may result in production inaccuracies. Exercises which target speech production should be complemented by prior perception training exercises. For speech perception, both discrimination and identification exercises should be introduced. Discrimination exercises, such as requiring students to discriminate between isolated sounds (e.g. /æ/ and /e/), can focus learners’ attention on articulatory and acoustic features of target phonemes. Identification exercises, on the other hand, can help learners associate a correct label to perceived acoustic signals, so when learners come across pronunciations such as [tʃi:p], they need to be trained to identify the length feature of the incoming signals and to associate the signals to the label cheap (rather than chip).

Awareness-raising production exercises, by the same token, should also embrace production of isolated sounds and production of words (and even phrases) in which target sounds are embedded. These go along the same line as discrimination and identification exercises in speech perception training, helping students to productively discriminating between confusable sound pairs and minimal word pairs.

Some example activities

In helping students differentiate between confusable L2 sounds, ESL teachers can introduce dictation exercises by asking students to listen to minimal pairs (e.g. ship, sheep), minimal sets (bee, tea, key, pea), or even sentences containing minimal pairs (e.g. Where are Jack’s bins and beans?) and to write down the words/sets/sentences that they hear. Students can also be engaged in production activities which require them to produce tongue twisters such as “Collecting the corrections is the role of the elderly” for /l/ and /r/, and “Woolen vests for wailing wolves are worn in the vast woodlands” for /w/ and /v/ (http://international.ouc.bc.ca/pronunciation/). Frequent practice of such tongue twisters can definitely help students approximate the correct model.
In alerting students to observable articulatory features, ESL teachers can make videos of their own production of certain sound contrasts (e.g. /ʃ/ and /s/) and compare the video clips with their students by focusing on the salient articulatory features such as the differences in the lip shape and the position of the jaw. Existing online resources, such as the BBC Learning English Website (http://www.bbc.co.uk/worldservice/learningenglish/multimedia/pron/index.shtml), which shows videos of speakers saying individual sounds, can also be used to arouse learners’ awareness of these features. More interactive activities which require students themselves to videotape their own production of the target sounds and which require other students to compare the video clips and judge the appropriateness of the articulatory features may also be introduced in class, so that students themselves can have a chance to become both models and evaluators.

Pronunciation learning is not a one-off or one-day process. For ESL students to learn a new L2 speech sound or to overcome a long-lasting or even fossilized pronunciation problem requires not just time and endeavour but also motivation and effective guidance. For ESL teachers to attain successful pronunciation teaching entails, on the other hand, not just enthusiasm and dedication and but also careful planning and systematic implementation. ESL teachers should not regard (remedial) pronunciation teaching as an additional component peripheral to their other teaching priorities and introduce L2 speech sounds sparingly on an ad hoc basis. The above guiding principles and suggested activities would lead to successful pronunciation teaching only if ESL teachers systematically and strategically follow the principles in their design and implementation of a comprehensive and research-informed pronunciation program.

**ESL pronunciation goals**

As discussed in the literature review, a standard native model (e.g. RP) has often been used as the norm for Cantonese ESL learners’ speech production and perception research. Such models have also often been adopted as the models for teaching by Cantonese ESL teachers at all primary, secondary and tertiary levels. While it is true that a standard native accent should be adopted by ESL teachers as the model (especially for
demonstration and discrimination purposes), it is debatable whether such a model should be taken as the ESL pronunciation goal and whether students should aim at achieving native-like accuracy.

The acceptability of an ESL learner’s pronunciation depends very much on the speech styles he/she is engaged in (casual vs. formal) and also on the receptiveness of the listener. In contexts where language accuracy is considered essential (e.g. the careful reading of a formal text) or when correct pronunciations are needed for the differentiation of words (e.g. minimal pairs), non-native pronunciations with noticeable deviations from the norms will be deemed unacceptable and will affect the listener’s understanding as well as his or her impression of the speaker’s English proficiency (Chan 2006a). Because of the possible adverse effects of faulty pronunciation, it is of course important that students be made aware of the need for correct pronunciation and be encouraged to achieve accuracy. However, certain English sounds, for example, /θ/ and /ð/, create so much difficulty to ESL learners that the inclusion of these sounds as pronunciation targets for English as an International Language (EIL) has been argued as inappropriate (Jenkins, 2002). Other pronunciation features, such as the dark [tʃ] articulation of the lateral /l/ in syllable-final positions, have also been disputed even in the native circle. In my opinion, the inclusion of the standard articulations of these sounds as ESL pronunciation targets is still necessary, especially given their high degree of error gravity. However, teachers should strike a balance between perfection and intelligibility and between cost-effectiveness and achievement of teaching goals. ESL teachers might want to give certain leeway to their students with respect to the achievement of native-like accuracy for such disputable L2 sounds.

Conclusion

In this article, I have outlined some principles and suggested some activities which are potentially beneficial to ESL pronunciation teaching based on documented speech production and perception problems encountered by Cantonese ESL learners in Hong Kong. It is argued that neither arena is to be disregarded, so both problems in production and problems in perception should shape the content and form the target of
teaching. Perceptual training is but one inevitable component of a holistic L2 pronunciation program going hand in hand with production training. While hiccups in production may cause stigmatization as a consequence of listeners’ poor impression of the speaker’s English proficiency, perception problems may also bring about misunderstanding and cause embarrassment. It is only through the training of both aspects that ESL learners be groomed towards successful mastery of L2 pronunciation. The principles given here are not meant to be innovative or groundbreaking. Instead, they may be so familiar to ESL professionals that they have been discounted. However, they are exactly the things that teachers should be constantly reminded of.

References


Conference on Arts and Humanities, Sheraton Waikiki Hotel, Honolulu, USA, 13-16 Jan 2005.


**Acknowledgements**

The work described in this article was supported by a competitive earmarked research grant from the Research Grants Council of the Hong Kong Special Administrative Region, China [Project No.: CityU 1455/05H]. The support of the Council is acknowledged.

**Biodata**

Alice Y.W. Chan is an Associate Professor at the Department of English, City University of Hong Kong. Her research interests include second language acquisition, English grammar, English phonetics and phonology, and lexicography.