

The Relationship between the Perception and Production of English Coda Clusters by EFL Thai Learners

Apichai Rungruang, Ph.D

Abstract—This study aims to investigate whether there is an interrelationship between the perception and production of 120 English final consonant clusters by native Thai speakers who learn English as a foreign language (EFL). The clusters were divided into three major groups: namely 64 two-member, 49 three-member, and 7 four-member clusters. 35 second-year native Thai university students took part in this study by taking two tests. An intelligibility test measured perception performance; a word-list reading test reflected their production performance. The framework to analyze the outputs was the markedness principle; the relationship between perception and production was examined by means of correlation analyses. The results revealed no relationship between their production and perception in the three types of final consonant clusters ($r = 0.18, 0.15, \text{ and } 0.06, N = 35, p > .05$). The findings also indicated that Thai participants did much better in the perception task than the production one, which covered 70 percent of all cluster tokens. Finally, the longest consonant margins or the four-member final consonant clusters posed the most difficulty perceiving and producing. This was compatible with the markedness principle in that the longer consonant margin, the more marked and difficult to perceive and produce.

Keywords—English coda clusters, markedness, principle speech production and perception.

I. INTRODUCTION

PERCEPTION and production in L2 interlanguage phonology have gained attention from many researchers [1],[2]. However, there is some controversy between the two. That is, some researchers state that both show interrelationship. Learners with good perception also have good production skill, poor perception skill corresponds with poor production skill. That means both skills develop interdependently [3], [4]. The next question is which skill develops first. In first language acquisition infants learn to perceive the sounds before being able to produce them. But there is no clear-cut agreement in L2 whether one precedes the other. Best claims that perception precedes production [5], [6]. Flege's Speech Learning Model states that speech perception influences speech production [7]. An increase in speech production performance is necessarily preceded by one in perception. Nonetheless, Sheldon and Strange strengthen the hypothesis that production may also precede perception in relation to L2 acquisition [8].

On the other side of the scale, some researchers claim that both perception and production have nothing to do with one another. In this study, the relationship between production and perception was investigated through English coda clusters by Thai learners. The focus of this study also was placed on the English coda clusters - all of which are not found in Thai to see how those had an impact on Thai learners' production and perception.

II. OBJECTIVES

The major objective of this study was to investigate the relationship between the students' production and perception in English coda clusters in the following two areas:

First, the relationship between speech perception and production in 120 English coda clusters to see the overall picture of student's performance;

Second, the relationship between speech perception and production in three sub-areas, namely two-member, three-member, and four-member codas which were regarded as marked structures for EFL Thai speakers.

III. HYPOTHESES

The researcher hypothesized that there would be a relationship between the students' production and perception in English coda clusters. In other words, if the students could identify the speech sounds, they should be able to produce them and vice versa. Another hypothesis was that students would show a stronger relationship between their speech perception and production in two-member coda clusters than the three and four-member clusters.

IV. THEORETICAL FRAMEWORK

This study employed the markedness principle to account for the findings. By and large, markedness is an abstract property of the no convention or unusualness, and difficulty of a sound. The unmarked elements are more basic, neutral, more universal, and first acquired; the marked elements are more specific, less frequent, and later acquired. It is claimed that all languages have CV syllable structure (C refers to a consonant; V refers to a vowel). In other words, CV is an unmarked form. So, any syllable structure that is more complex than the CV one is regarded as a marked structure. To be more precise, if the number of consonants (or the length of onset/coda) increases, the level of markedness

increases. The longer consonant margin, the more marked and difficult to perceive and produce. For instance, CCV is more marked than CV; CCCV is more marked than CV and CCV. By the same token, VC is less marked than VCC, VCCC, etc. To put this notion into action, the researcher chose Markedness Differential Hypothesis (MDH) by Ekman [9] to account for all perceptive and productive difficulty or ease from Thai speakers who learned English as a foreign language (EFL learners). Below are the MDH's claims.

1) Those target language (TL) structures that are both different and relatively more marked than the corresponding structures in the native language (NL) are predicted to be more difficult.

2) The degree of difficulty will correspond directly to the degree of markedness.

3) Those TL differences that are not more marked will not be more difficult.

Let's consider English and Thai syllable structures (more detail will be discussed in the next part). While Thai has only simplex codas, English has many complex coda clusters. Therefore, native English speakers tend to have less difficulty learning Thai coda than native Thai speakers who learn English. In conclusion, markedness is an abstract property of the no convention or unusualness, and difficulty of a sound. The unmarked elements are more basic, neutral, more universal, and first acquired; the marked elements are more specific, less frequent, and later acquired. The length of margin, and marked/unmarked sounds in the mother tongue or NL and TL play a significant role in phonological acquisition.

V. THAI AND ENGLISH FINAL CONSONANT STRUCTURE

In Thai, only certain simplex codas are found. To be more specific, Thai has the following eight sounds in final position: voiceless stops /πτκ/, voiced nasals /μνN/, and semivowels /φω/. In other words, coda clusters are not allowed. Unlike Thai, English has much more complex codas. The maximum number of final consonant in English is 4. However, Shockey [10] claims that in reality it is so common for native English speakers to reduce four-consonantal codas to tri-consonantal ones. Below are some of them.

- CC			
gulp	/-λπ/	jobs	/-βζ/
- CCC			
sculpt	/-λπτ/	world	/-♦λδ/
- CCCC			
thousandths	/-νδTσ/	texts	/-κστσ/

Fig. 1 Samples of consonant structure

As mentioned above, Thai syllable structure is limited to single-consonant codas, but English has much more complex structure than Thai. Therefore, it is very possible that Thai speakers tend to have some difficulties pronouncing and perceiving English coda clusters. It was hypothesized that Thai speakers would perform better in two-member

consonants than three and four-member ones. In the present study, the total number of tokens was 120. Tokens in -CC, -CCC, and -CCCC were 64, 49, and 7, respectively.

VI. RESEARCH METHODOLOGY

A. Participants

The study was carried out in July to August 2012. 35 second year English major students participated in this study. All of them took two required English courses (Fundamental English and Developmental English) and a basic writing course when they were in their first year of BA study. They had learned English for at least 10 years and usually used English only in class. Their mother tongue, Thai, was used outside the classroom and in everyday conversation. None of them had studied abroad, nor spent extended periods of time in English speaking countries.

B. Research Instruments

There were two major tools: an intelligibility test and a word-list reading test. The intelligibility test was to examine how well the participants were able to identify coda clusters. The researcher made 120 items from selected 120 coda clusters and also made four choices (a, b, c, and d) for each item. Before the test was carried out, the sounds from a native speaker of English was recorded. The researcher asked an American native speaker to pronounce each word twice. The pause between each word was approximately 4-5 seconds. The native speaker reading was recorded in a sound-proof room through a phonetic computer software named 'Praat'. In this test, the participants were asked to indicate which English word they had just heard.

The second tool was a word-list reading test. Since this study also examined participant's production, all 120 coda clusters from the intelligibility test were listed on a sheet. Unlike the first tool, the second tool did not have any choices. Each participant was tape-recorded in a face-to-face manner with the researcher. What the participants had to do was to read out words twice from items 1 to 120. All their speech sounds were recorded by a phonetic computer software 'Praat'.

C. Data Collection Procedure

The researcher ran the intelligibility test first by scheduling all 35 participants to sit in a sound-proof room. They were not informed of the real purpose of the study but were instead told to listen to the record of 120 words as a part of some research. Each word was read twice. A slight pause between words marked the end of the preceding word. They listened carefully to each word, and circled the best choice (a, b, c, and d). It took them 5 minutes to complete this task.

To ensure that the participants did not have a clue what the researcher would do to them in the next task, the researcher waited 4 weeks before recoding individual participants' speech sounds. They had to pronounce 120 target words. Individual participants used a microphone to pronounce each word. During the tape recording the researcher raised a finger as a signal to have the participant pronounce the next word. It took approximately 18 minutes to record all 120

words for each participant. Below are all the tokens used in this study.

D. Research Validity And Reliability

To establish content validity in the intelligibility test, all tokens were collected from different textbooks and previous studies, and then had them checked by three phonetics course instructors. Two of the experts received a doctoral degree and one held a master's degree. All tokens reflected possible types of English complex onsets. Another measure used to set up content validity was through two English native speaker raters, both of whom hold Bachelor degrees from accredited universities in the United States and Australia. Before doing the rating, the two raters were trained to understand what the study aimed to investigate and how to investigate them. Both raters independently rated the participants' speech sounds. The other measure to establish validity is the researcher assessed the appropriateness of the test by running a pilot test with 10 students. Based on the pilot test outcomes, a few changes were made to the test. In its final form, the test was printed on two (double-sided) A 4 pages; the average time to complete the test was 5 minutes.

In terms of reliability, to ensure that two raters agreed on their judgment or to confirm that the coding was reliable, interrater correlations were calculated through Pearson Product Moment Correlation (r). A computer software was operated to find the inter-rater reliability score (r). It turned out that the r value was 0.933, which referred to a very strong relationship between the two raters [10].

VII. FINDINGS AND DISCUSSION

A. Overall Picture

After the 35 subjects completed an intelligibility test to measure their perception performance and a word-list reading test to assess their production performance, it turns out that they did much better in the intelligibility test. That is, from the total score of 120 (from 120 tokens), the average scores in perception and production were 83.74 (70%) and 63.11(53%), respectively.

TABLE I
RESEARCH RESULT

	Types of performance	Mean (mean % correct)	Std. Deviation	r	p-value	No of Tokens
Overall (2-3-4 member consonants)	Perception	83.74 (70%)	10.74	.20	.24	120
	Production	63.11 (53%)	20.87			
Two-member	Perception	46.71 (73%)	6.60	.18	.29	64
	Production	37.61 (59%)	11.68			
Three-member	Perception	33.31 (68%)	4.68	.15	.36	49
	Production	23.50 (48%)	8.86			
Four-member	Perception	3.71 (53%)	1.20	.06	.73	7

	Types of performance	Mean (mean % correct)	Std. Deviation	r	p-value	No of Tokens
Overall (2-3-4)	Perception	83.74 (70%)	10.74	.20	.24	120
	Production	2.00 (29%)	1.43			

Thus, students performed much better in the perception test than the production one. Not surprisingly, when Pearson correlation coefficients (r) was calculated to see the relationship between both performances, it showed no significant difference (the r value was 0.20). That means if students could identify the speech sounds, they might or might not be able to produce them and vice versa. Besides the difference of mean scores between the two performances, it probably did not reach significance due to the large standard deviation (SD) which reflected the amount of score variation or dispersion from the average mean. Students showed better score variation in perception than production (10.74 and 20.87, respectively). In principle, the lower SD, the better it is. An interesting question is posed: why did the Thai subjects perform better in the perception task than the production task? It is very possible that production performance requires a much more great deal of practice and dedication than perception performance. When speech sounds are produced, they are modified through active articulators (lower lips, lower teeth, and tongue) and passive ones (upper lip, upper teeth, alveolar ridge, hard and soft palate). Before turned into a speech sound, the airflow is modified through the place of articulation or where the speech sounds are changed and the manner of articulation or how speech sounds are produced to a particular speech sound correctly. To reach the level of speech sound accuracy, students need a great deal of practice and understand how airstream mechanism work. Therefore, the first hypothesis claiming that the relationship between the students' production and perception in English coda clusters existed was rejected. Let's narrow it down to individual type of coda clusters from two-member, three-member, and four-member codas, respectively.

B. Two-Member Codas

TABLE I shows that the mean score in perception was higher than that in production (46.71 or 73% and 37.61 or 59%, respectively). Notice that students performed two-member clusters better than in the overall member clusters. This is not surprising since the two-member clusters is in the lowest level of marked structure scale among the three types for Thai speakers whose native language do not have complex final consonants. However, some two-member tokens posed a challenge to the Thai speakers as well (this will be discussed later). Like mean scores, SD in perception less dispersed from the mean score than that in production (6.60 and 11.68, respectively). In terms of the relationship between their perception and production, the r value was 0.18, which reflected a weak relationship between the two tasks. The p -value was 0.29, which referred to no statistical significance.

C. Three-member codas

49 three-member codas were examined. The mean score in perception was still higher than that in production (33.31 and 23.50, respectively). Like mean scores, SD in perception was less dispersed from the mean score than that in production (4.68 and 8.86, respectively). However, compared to the two-member codas, the three-member ones showed better SD. Students seemed to perceive and produce speech sounds close to the mean score. To see the relationship between the two tasks, the r value was calculated (0.15), which reflected a weak relationship.

D. Four-member codas

7 four-member codas were tested. Compared to the previous two types of codas, the mean scores in perception was still higher than that in production (3.71 and 2, respectively). Like mean scores, SD in perception was slightly less dispersed from the mean score than that in production (1.20 and 1.43, respectively). However, when compared to the two-member codas, the three-member ones showed better SD. Once again, the findings also revealed no significant difference between two performances (the r value was 0.06).

Since no significant difference was found in three types of coda clusters, this brings us to the second hypothesis stating that students would show a stronger relationship between their speech perception and production in two-member coda clusters than the three and four-member clusters. In light of the data obtained, this hypothesis was confirmed. However, the r value in the two-member codas was slightly higher than the three-member ones (0.18 and 0.15, respectively). But it revealed a rather obvious difference between two-member and four-member codas (0.18 and 0.06, respectively).

The findings also showed that the subjects had some difficulty perceiving and producing the sounds that contrast much more from the sounds of their native phonological inventory. The notion of perceiving sounds between nonnative and native sounds plays a crucial role accounting for the subject's performance. MDH's claims, as a major theoretical framework in this study, provide a good reason for this. That is, the ease or difficulty of a nonnative contrast depends on the similarity to the corresponding sounds of the native phonology. In principle, L2 learners tend to do well if the particular sounds in L2 are very close to those in their native language. Based on the current results, Thai students, as mentioned earlier, do not have coda clusters in their native language. They did poorer when they faced long consonant margins, particularly the four-member codas. Therefore, the difficulty predicted on the basis of markedness principle was fully supported. The more marked (the long margin) pose the most difficulty for Thai speakers. Interestingly, when paying close attention to some of the tokens in the two-three- and four-member codas, the researcher found that most of the very low score coda clusters were related to English plural and past forms. Here are some examples.

Two-member codas

march/- $\diamond\Sigma$ /(3/19)
breathed/- $\Delta\delta$ /(25/6)
clothes/- $T\sigma$ /(5/20)
begged/- $\gamma\delta$ /(8/16)

Three-member codas

depths /- $\pi T\sigma$ /(7/10)
midst /- $\delta\sigma\tau$ /(7/3)
sevenths /- $\nu T\sigma$ /(6/10)

Four-member codas

texts /- $\kappa\sigma\tau\sigma$ /(16/2)
thousandths/- $\nu\delta T\sigma$ /(10/22)

Fig. 2 Samples of low score tokens in each cluster.

Phonetic symbols are in slanted brackets. In parenthesis, the first number refers to perception score; the other refers to the production score. For example, in the term 'march' /- $\diamond\Sigma$ /(3/19) only 3 subjects (out of 35) were able to perceive it correctly and 19 subjects (out of 35) could pronounce it accurately. It is very possible that the low use of regular past and plural form by Thai speakers reflects a strong L1 constraint against final consonant clusters. These findings were compatible with what Lardiere [11] found when she studied native speakers of Mandarin and Hokkien who used English as a second language. Without coda clusters in their Chinese phonological system, her subjects faced the same situation and had some difficulties producing and perceiving the English regular past. Besides English plural and past forms, some other terms posed a tremendous challenge to the Thai speakers. They are: 'march' and 'midst'. For the former term, students were confused with the term "marsh". They perceived them as the same term, but some were able to pronounce it correctly. The worst score went to the latter one. In the term "midst", not many Thai speakers could perceive and produce it correctly. Only three students could pronounce this word accurately. They dropped a fricative /s/ since their phonological inventory did not have this sound.

VIII. CONCLUSION

This research has presented some preliminary evidence on the relationship between speech perception and production. It was conducted when the subjects were second-year students. This evidence shows that there was no correlation between perception and production. However, it is too soon to claim that a relationship between perception and production does not exist. A follow-up study will be carried out when the students, as advanced learners, are in the final year of their study. Therefore, a complete picture will be drawn and will reflect changes after they have gone through a number of English courses, namely phonetics, speaking, or conversation courses, etc. As a result, it will be very interesting to find out how much change they have made in their speech perception and production performances and whether the relationship between speech perception and production truly exists.

ACKNOWLEDGMENT

This study was financially supported by a research grant No R2557C105 from Division of Research Administration, Naresuan University, Thailand. I also would like to thank all students who volunteered to participate in and take perception

and production tests in this study.

REFERENCES

- [1] De Wilde, E. 2009. *Perception and production in second language phonology. The effect of audiovisual training on the acquisition of the English dental fricatives*. MA thesis, Universiteit Gent. Netherlands.
- [2] Peperkamp, S and Bouchon, C. 2011. The relation between perception and production in L2 phonological processing. *INTERSPEECH 2011, 12th Annual Conference of the International Speech Communication Association*, Florence, Italy, August 27-31, 2011. ISCA 2011. pp.161-165.
- [3] Bradlow, R A., Pisoni, B. D., Akahane-Yamada, R. & Tohkura, Y. 1997. Training Japanese listeners to identify English /r/-/l/: IV. Some effects of perceptual learning on speech production. *Journal of Acoustic Society of America*, 101, pp. 2299-2310.
- [4] de Jonge, E. C. 1995. *Interlanguage phonology: Perception and production*. Unpublished doctoral dissertation, Indiana University
- [5] Best, C. T. 1995. A direct realist perspective on cross-language speech perception. In W. Strange (ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research*, pp. 167-200. Timonium, MD: York press.
- [6] Best, C. T, Gerald W. McRoberts, Elizabeth Goodell. 2001. Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109 (2), pp. 775-794.
- [7] Flege, J. E. 1995. Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech production and linguistic experience: Issues in cross-language research* pp. 233- 277. Timonium, Md: York Press.
- [8] Sheldon, A and Streange, W. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3, pp.243-261.
- [9] Eckman, F. R. 1977. Markedness and the contrastive analysis hypothesis. *Language Learning*, 27, pp. 315 - 30.
- [10] Shockey, L. 2003. *Sound patterns of spoken English*. Oxford: Blackwell.
- [11] Lardiere, D. 2003. Second language knowledge of [+/-Past] and [+/-Finite]. In Juana M. Liceras et al. (Eds.), *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002)* pp. 176-189. Somerville, MA: Cascadilla Proceedings Project.