Coda Discrimination is Governed by Acoustic and Phonological Constraints: A Case Study of the Generalized Linear Model*

Chu, Man-ni**
Department of Applied English
Der Lin Institute of Technology

Lin, Hui-wen
Biostatistics Research and Consulting Center
Taipei Medical University

ABSTRACT

A series of studies has shown that consonants can be categorized by their acoustic character. A gating experiment was designed to recognize the factors used to identify terminal plosives correctly. Nine native residents of Shantou, in China's Guangdong Province, were recruited as subjects. The dialect spoken in Shantou does not allow */ap/ and */at/. Participants were asked to identify /taC/ and /kaC/ under conditions where C is one of the four choices /p, t, k, ʔ/. The overall rate of accuracy, from highest to lowest, was /p/ > /k/, /ʔ/ > /t/. In general, the similarities in the formant transitions of /t, k, ʔ/, also observed by Tartter et al. (1983), confused subjects, which resulted in a higher recognition rate for the labial stop. Results suggested that the subjects' language background helped discriminate the remaining sounds; for example, /k, ʔ/ were correctly identified more often than /t/.

Key Words: Shantou, terminal plosives, accuracy rate, second formant transitions

---

* Heartfelt thanks to two anonymous reviewers for insightful comments and criticisms on earlier drafts of this paper. It goes without saying that any errors or infelicities that remain are ours.

** The corresponding author's email address: mannichu@gmail.com

1. Introduction

The acoustic nature of consonants has been described by previous researchers (Stevens 1989; Stevens and Blumstein 1978, 1981). In general, the relationship between consonants and their adjacent vowels has been described as follows. Ladefoged (1975) reported that the formants of the vowel (e.g., F2 and F3) are crucial in discriminating the place of adjacent stops. In short, F2 and F3 decrease when the vowel is adjunct to /p/; they are marked as 'velar pinch' in /k/; however, F2 and F3 are fairly steady when they occur before or after /t/. Although no specific description of the interaction between /ʔ/ and its adjacent vowel has been given, /ʔ/'s frequently occur as an allophone of /t/, which presumably makes us guess that its formant transition is similar to /t/. Because one phoneme may be in-various in terms of its acoustic nature, variability was shown in the behavior of phone or multi-phoneme (Klatt 1980; Ohala 1992). In addition, Ohala et al. (1994) claimed that the identification of a sound is influenced by its neighboring sounds. In that study, subjects were asked to restore the vowel when its adjacent consonants /b/ or /d/ were masked by VCə utterances. Results indicated that the neighboring context did affect subject vowel identification. Thus, in this study a perceptual experiment was conducted to examine the ability of individuals to discriminate between the four final stops.

Lin and Chen (1996) reported that Shantou, a local dialect spoken by some 370,000 people in Guangdong Province, China, has three codas, /p/, /k/ and /ʔ/. Therefore, the /t/ identification was presumed to be the coda that would be least recognized correctly, as participants could be expected to miss sounds that do not exist in their own system. van Wieringen (1995) reported that native speakers of Japanese, whose language does not allow coda consonants, face greater difficulty in identifying terminal plosives as compared to native speakers of Dutch, whose language allows the CVC structure. Tsukada et al. (2007) found that subjects of different language backgrounds react differently in identifying word-final stops. In Tsukada's study, stimuli were drawn from American English, varying between released and non-released sounds, and Thai, with only unreleased terminal plosives. To Cantonese, Korean, and Vietnamese subjects, who always allow unreleased terminal stops, English pairs are more easily recognized than Thai pairs. The conclusion is that familiarity with a particular sound in a foreign language affects non-speaker abilities to discriminate and identify sounds correctly. In pair discriminations, /t/ and /k/ are the least easy to discriminate from one another. This means that /t/ and /k/ are perceptually close.

The gating experiment gives us information regarding the process by which individuals perceive sounds because subjects are presented with information gradually as they perceive a sound at each gate. Ciocca et al. (1994:1134) mentioned that F2 and F3 can differ /p, t, k/ in the middle of a vowel, inferring an assumption of the experiment that subjects perceive terminal stops in the middle of vowel formants.

In the present experiment, we attempted to discover the factors which dominate in subject identification by identifying which fragment causes subjects to detect coda characters. Specifically, we examined accuracy rates and tried from such observations to explain coda identification from both the acoustic and phonological aspects.

2. Methodology

2.1 Materials and Procedure

In order to examine the four sets of coda,[1] stimuli were recorded by a male native speaker of Tainan Southern-Min, which has /p/, /t/, /k/, and /ʔ/, using CoolEdit Pro 2.0 with a 44.1 kHz sampling rate and a 16 bit quantization level. The experiment presented six H-tone tokens of CVC wave forms, which varied in the extent to which they were cut short. This was done in 20ms steps from the end, such that the longest was intact and the shortest ended 100ms earlier. The start point was set at the beginning of the onset stop, where the initial closure was included. Onset C varied between /t/ and /k/, so that we had 6(steps)×2(C-)×4(-C)=48 stimuli (see Appendix A). Shantou is a non-alphabetic language that uses Chinese characters to represent its lexicon. In order to prevent subjects from misrecognizing terminal stops, subjects were first trained to gain an appropriate level of familiarity with IPA. Subjects were asked to indicate which terminal stop was located in a series of sounds. A 24-stimuli pretest containing the four choices [p, t, k, ʔ] was run to make sure the subjects knew what they would be asked to do during the formal experiment. Subjects who passed the pretest may participate in the next experiment. Then, a 48-stimuli identification test presented subjects with a total of four choices [p, t, k, ʔ]. In order to focus their

1. It has been reported that different languages may have different phonetic features in producing the same consonant; however, since Tainan Southern-Min and Shantou belong to the same Southern-Min system and it is difficult for Shantou natives to produce the final /-t/ which is non-existent in their language, we believe that the languages difference between the listener and the speaker can be ignored.

attention, all the sounds were immediately preceded by a warning signal. All the sounds, presented in a random order determined by the computer software, appeared on the computer screen for subjects to select the one choice that best matched the most recent consonant heard. Each consonant sound was presented three times, giving a total of $48 \times 3(\text{times}) = 144$ stimuli. In each section, the 48 stimuli were tested randomly and subjects needed to attend to the same experiment three times. Subjects were required to guess what they heard if they were not sure. The whole experiment took about 30 minutes. There were two options for the data-calculation. In the first option, all the answers in the raw data were counted. In the second, an answer was counted only when a subject consistently gave two same answers out of three experiments. We adopted the second method for stronger reliability, on the one hand, and for simplification, on the other hand. That is, if all the responses were collected, another fixed random effect model should be proposed since the three answers were highly correlated. In order to simplify the model and provide evidence to show that the logistic regression model is adequate to analyze the data, the same two answers were counted as the answer to one particular stimulus. If none of the answers were the same, the answer would be classified as 'undecided'.

2.2 Participants

A total of 18 native speakers of the Shantou dialect of Chinese who self-reported having no hearing ϕ problems participated in the experiment. Only the data of nine subjects were used for the experiment because they passed the threshold of 10/24 in the pretest. Evidence indicates that the probability that the subjects had made a guess by passing the threshold is 0.03, $p < 0.05$. That means that those 9 (4 males, 5 females, aged 9-52, Mean 33, SD 15) participants who passed the threshold by guessing the answer correctly 10 out of 24 is less than .05.

2.3 Statistical Method

Among classical linear models (e.g., ANOVA, t-test), we usually assume independent outcomes that are normally distributed with equal variance. These assumptions are manifestly uncertain when the outcome variable is, for example, the correctness or incorrectness of a test, the success or failure of an operation, etc. The characteristics of such cases mean that: (a) normality is not a reasonable assumption, (b) a linear model is usually not appropriate, and (c) variance is generally dependent upon the mean. It is natural to model an outcome variable as binary or dichotomous with probability (Hosmer and Lemeshow 1989), so we propose a logistic regression model (Hosmer and Lemeshow 1989;

Agresti 1995) to model our data.

The designed logistic regression model is:

$$\log\frac{\Pr(Y=1|coda,onset,gate)}{\Pr(Y=0|coda,onset,gate)}=\beta_0+\beta_{1j_1}coda+\beta_{2j_2}onset+\beta_{3j_3}gate$$

Where the fixed variables coda, onset, and gate represent categorical data and Y is the outcome variable. Y=1 represents a correct result; Y=0 represents an incorrect result. Parameter β is a vector, which can be estimated using Maximum Likelihood Estimates (MLEs). The interpretation of the β_j parameter estimates is as the additive effect on the log **odds ratio** (log **odds ratio** is a log ratio of the percentage correct and percentage of incorrect.) for a unit change in the jth explanatory variable. In the case of a dichotomous explanatory variable, for instance, coda, e^{β} is the estimate of the odds ratio of having the outcome for, say, /p/, /k/, and /ʔ/ compared with /t/. The model has an equivalent formulation

$$P=\frac{e^{\beta_0+\beta_{1j_1}coda+\beta_{2j_2}onset+\beta_{3j_3}gate}}{1+e^{\beta_0+\beta_{1j_1}coda+\beta_{2j_2}onset+\beta_{3j_3}gate}}$$

P is percentage correct given in coda, onset and gate condition.

2.3.1 The Stepwise Method of Selecting Variables (Cohen 1991)

As we were also interested in determining which variables have the most significant effect on the results, we used a forward stepwise method, i.e., the Wald statistic for variable selection. The Wald statistic was calculated for variables in the model to determine whether a variable should be removed.

Where β is the vector of maximum likelihood estimates associated with the J-1 dummy variables and **V** the asymptotic covariance matrix for β:

the Wald statistic is $Wald=\hat{\beta}V^{-1}\hat{\beta}$

The asymptotic distribution of the Wald statistic is chi-square, with degrees of freedom equal to the number of parameters estimated.

2.3.2 Forward Stepwise (FSTEP)

FSTEP Algorithms obtain a final model by starting with all the variables suspected to affect the coda identification ability of subjects. To determine the subset of suspect variables, all the variables are estimated as a function of parameters and likelihood functions. In examining variable validity, we obtain the necessary information by checking the MLE parameters, the prediction probabilities, and the likelihood function for the current model. A variable is removed from consideration when it is determined to be insignificant in the model as calculated by the Wald statistic. Then, the interactive variables are added to result in a model containing the new variable. The new model is then run and re-evaluated.

3. Results

Logistic regression was performed to examine which factors affect the coda identification of subjects in terms of the percentage of correct answers given. The method (forward: wald) was chosen. Table 1 represents the processing in step 1, where the coda is the variable which enters first; this means that it affects the percentage correct most. Then, the gate level entered in step 2, which is esti-

Table 1 The Steps of the FSTEP

Variables in the Equation

		?	S.E.	Wald	df	Sig.	Exp(?)
Step 1(a)	coda			13.484	3	**.004**	
	coda(1)	1.157	.325	12.683	1	.000	3.182
	coda(2)	.511	.341	2.249	1	.134	1.667
	coda(3)	.560	.339	2.726	1	.099	1.750
	Constant	−1.609	.258	38.854	1	.000	.200
Step 2(b)	coda			13.932	3	.003	
	coda(1)	1.200	.331	13.128	1	.000	3.320
	coda(2)	.527	.346	2.320	1	.128	1.694
	coda(3)	.578	.344	2.814	1	.093	1.782
	gatelevel			14.576	5	.012	
	gatelevel(1)	−.811	.374	4.692	1	.030	.445
	gatelevel(2)	−1.159	.396	8.553	1	.003	.314
	gatelevel(3)	−1.066	.390	7.476	1	.006	.344
	gatelevel(4)	−.811	.374	4.692	1	.030	.445
	gatelevel(5)	−.247	.352	.493	1	.482	.781
	Constant	−.987	.335	8.664	1	.003	.373

mated to be the second important factor to affect the percentage correct. The onset was not included in the model because it did not significantly affect listeners' performance. Thus, the results in Table 2 illustrate that two variables are included in the final model. That is, 'coda' and 'gate' affected coda identification, but 'onset' did not. In short, the percentage of each correct differed significantly from each other while the input (coda) and the gate varied. The final model contains 'coda' and 'gate', where both were determined to have significance after the modified model parameters were estimated.

The results shown in Table 2 indicate that in terms of percentage correct, coda and gate were significant (both $p<.05$). Specifically, the percentage correct of /t/ versus /p/ was significant ($p<.05$) but there was no significant difference between /t/ and /k/ and /t/ and /ʔ/ (both $p>.05$). Compared to /t/, the percentage corrects of /p/, /k/ and /ʔ/ were 3.32, 1.69, and 1.78 times that of /t/, respectively. That means /p/ was the easiest of the four consonants to recognize correctly. In order to resolve the suspicion that /p/ might be used by subjects as a 'default' answer, a t-test was performed, with the result indicating that /p/ was not a 'default' because the response distribution was still within 95%. In other words, the subjects did not respond by choosing /p/ more often than the other consonants, regardless of the number of inputs. In addition, gate 6 (original duration) did not differ significantly from gate 5 (20 ms shorter from the end in dura-

Table 2 The Final FSTEP Model

		β	S.E.	Wald	df	Sig.	Exp(β)
Final model	coda			13.932	3	**.003**	
	coda(1)	1.200	.331	13.128	1	.000	3.320
	coda(2)	.527	.346	2.320	1	.128	1.694
	coda(3)	.578	.344	2.814	1	.093	1.782
	gatelevel			14.576	5	**.012**	
	gatelevel(1)	−.811	.374	4.692	1	.030	.445
	gatelevel(2)	−1.159	.396	8.553	1	.003	.314
	gatelevel(3)	−1.066	.390	7.476	1	.006	.344
	gatelevel(4)	−.811	.374	4.692	1	.030	.445
	gatelevel(5)	−.247	.352	.493	1	.482	.781
	Constant	−.987	.335	8.664	1	.003	.373

tion) in terms of percentage correct, and there was a significant difference between gate 6 and gates 1, 2, 3, and 4, respectively (all $p<.05$), where gate 1 represents the very first 20 ms from the start point of the onset. In other words, the subjects were able to detect the coda in the middle of the previous vowel. Specifically, compared to gate 6, the percentage correct of gates 1, 2, 3, 4, and 5 were .45, .31, .34, .45, and .78 times that of gate 6, respectively. In short, the gate and the coda are two important factors that affect subject coda identification. The coda /p/ seems to be correctly recognized most often and the identification rate in decreasing order is /p/>/k/, /ʔ/>/t/. Furthermore, the identification rates in gates 1, 2, 3, and 4 were significantly different from gate 5 and gate 6 since more information was provided.

Table 3 The Results of Estimated Marginal Means: coda

Pairwise Comparisons

(I) coda	(J) coda	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
p	k	.14[a]	.064	1	.025	.02	.27
	h	.13[a]	.063	1	.036	.01	.26
	t	.23[a]	.057	1	.000	.11	.34
k	p	−.14[a]	.064	1	.025	−.27	−.02
	h	.00	.061	1	.877	−.13	.11
	t	.08	.055	1	.127	−.02	.19
h	p	−.13[a]	.063	1	.036	−.26	.00
	k	.01	.061	1	.877	−.11	.13
	t	.09	.055	1	.089	−.01	.20
t	p	−.23[a]	.057	1	.000	−.34	−.11
	k	−.08	.055	1	.127	−.19	.02
	h	−.09	.055	1	.089	−.20	.01

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable correct

a. The mean difference is significant at the .05 level.

Table 4 The Results of Estimated Marginal Means: gate

Pairwise Comparisons

(I) gate level	(J) gate level	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
gate1	gate2	.06	.067	1	.409	−.08	.19
	gate3	.04	.068	1	.541	−.09	.18
	gate4	.00	.070	1	1.000	−.14	.14
	gate5	−.11	.074	1	.125	−.26	.03
	gate6	−.17[a]	.079	1	.030	−.33	−.02
gate2	gate1	−.06	.067	1	.409	−.19	.08
	gate3	−.01	.064	1	.831	−.14	.11
	gate4	−.06	.066	1	.401	−.18	.07
	gate5	−.17[a]	.071	1	.017	−.31	−.03
	gate6	−.23[a]	.076	1	.003	−.37	−.08
gate3	gate1	−.04	.068	1	.541	−.18	.09
	gate2	.01	.064	1	.831	−.11	.14
	gate4	−.04	.067	1	.534	−.17	.09
	gate5	−.16[a]	.072	1	.030	−.30	−.01
	gate6	−.21[a]	.076	1	.005	−.36	−.06
gate4	gate1	.00	.070	1	1.000	−.14	.14
	gate2	.06	.066	1	.401	−.07	.18
	gate3	.04	.067	1	.534	−.09	.17
	gate5	−.11	.073	1	.119	−.26	.03
	gate6	−.17[a]	.078	1	.027	−.32	−.02
gate5	gate1	.11	.074	1	.125	−.03	.26
	gate2	.17[a]	.071	1	.017	.03	.31
	gate3	.16[a]	.072	1	.030	.01	.30
	gate4	.11	.073	1	.119	−.03	.26
	gate6	−.06	.081	1	.480	−.22	.10
gate6	gate1	.17[a]	.079	1	.030	.02	.33
	gate2	.23[a]	.076	1	.003	.08	.37
	gate3	.21[a]	.076	1	.005	.06	.36
	gate4	.17[a]	.078	1	.027	.02	.32
	gate5	.06	.081	1	.480	−.10	.22

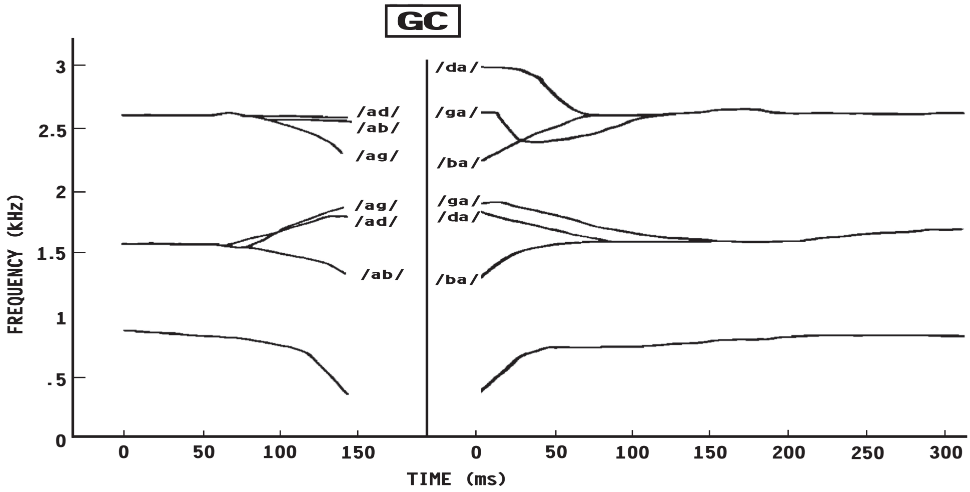
Pairwise comparisons of estimated marginal means based on the original scale of dependent variable correct

In addition, Estimated Marginal Means was run to examine the pairwise comparisons shown in Table 3 and Table 4. Results indicated that the percentage correct of /p/ was significantly better than /k/, /ʔ/, and /t/ (all $p<.05$), and there was no significant difference among the three (all $p>.05$). The percentage correct of gate 6 was significantly different from that of gates 1, 2, 3, and 4 and there was no significant difference between that of gate 6 and gate 5. The percentage correct of gate 5 was significantly different from that of gates 2 and 3 (both $p<.05$) and there was no significant difference between that of gate 5 and that of gates 1 and 4 respectively (both $p>.05$). There was no significant difference among the percentage correct of gate 1, 2, and 3 (all $p>.05$).

4. Discussion

Terminal plosives in American English are more easily identified when they are released (Wang 1959:71). In other words, final unreleased terminal stops are relatively likely to cause errors (Malecot 1958:376). In addition, Redford and Diehl (1999) found that onset formant transitions were more easily identified than those of codas because onset stops always contain at least a burst feature (released). Nevertheless, Chinese terminal stops are similar to Thai in that they contain no release information, which makes hearing them contingent on other information and formant transitions. A series of discussions has previously taken place about dispersion phonology. Mielke (2002) examined /h/ perception in English, French, Turkish, and Arabic, and stated the lack of robust acoustic cues leading to perceptual similarity, which produces to the lack or the loss of contrast. On the other hand, Lindblom and Engstrand (1989) observed that the tendency is to maximize perceptual distinctiveness. In other words, speech sounds need to be contrasted as much as possible in order to separate sounds in perceptual space. An experiment conducted by Tartter et al. (1983) noted similarities between /ad/ and /ag/ (see the left side of Figure 1). Figure 1 indicates that stops in both onset and coda positions reveal similar acoustic patterns: the second formant transition of /g/ and /d/ after /a/ goes up, where that of /b/ goes down, resulting in a significant difference in perception. As /d/ and /g/ are both voiced, we believe that there is no difference between voiced and voiceless stops in terms of acoustic character. In addition, Figure 2 illustrates the stimuli /tap/, /tat/, /tak/, and /taʔ/ in the study. There was a slight difference between /p/ and the other codas regarding the second formant. Thus, the conclusion is that the place of articulation F2 in vowel transition enhances distinctive stops, which distinguishes the essential combining feature, turning out to maximize

Figure 1 Coda Locus Effect in Comparison with the Onset (from Tartter et al. 1983:717)



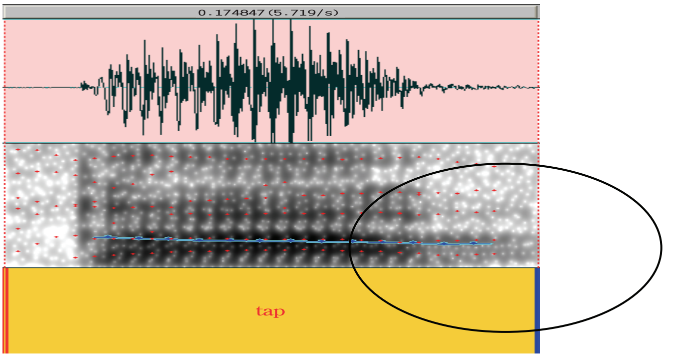
perceptual contrasts.

It was reported that subjects detect the consonant in the middle of the previous vowel (Ciocca et al. 1994 and supported in Figure 1). The F2 difference between /ag/, /ad/, and /ab/ lies around 60 ms shorter from the end of the sound. In other words, /ad/ and /ag/ have slopes that are similar in F2, which makes a significant difference between /p/ and /t, k/ after /a/. On the other hand, the formant transition of /ʔ/ is presumed to be similar to /t/, since it has been argued that /ʔ/'s frequently occur as an allophone of /t/. To summarize, the second formant transition is the main effect causing /p/ to differ from /t, k, ʔ/. Furthermore, the way to calculate the answer makes the results more reliable, where subjects choosing two different answers of the same input are considered 'undecided', i.e., missing data. That is to say, whenever the answer was labeled /p, t, k, ʔ/, subjects chose /p, t, k, ʔ/ twice, which is one way to assure that subjects really were confident about what they heard. They chose what they believed that they perceived, which is further evidence for enlarging the distinction between /t, k, ʔ/ and /p/.

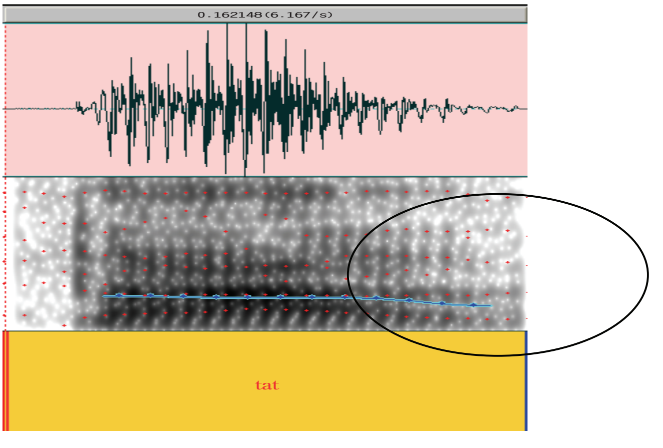
On the one hand, we assumed that /p/ was identified more correctly than the other codas by subjects and, overall, that subjects could not detect codas until they terminate 40ms prior to the end. Previous studies observed that if the word (CaC) combination does not exist in the subject's lexicon, it would be difficult for them to identify those non-existent words. Table 5 represents the CaC combinations of Shantou and Tainan (stimuli) dialects, showing that Shantou does not

Figure 2 The Stimuli of /tap/, /tat/, /tak/ and /taʔ/

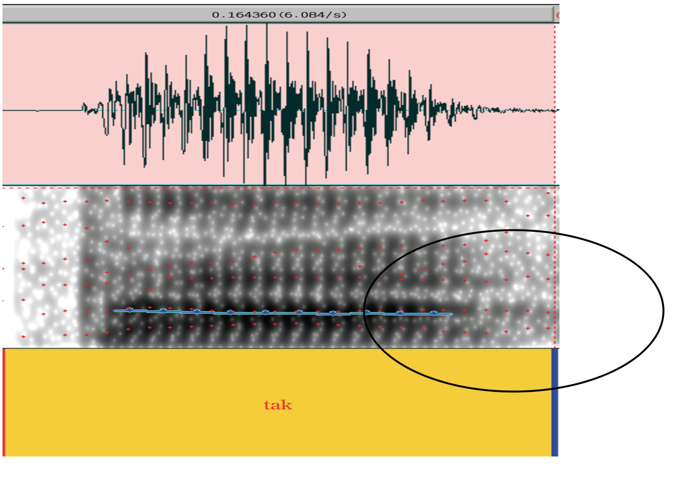
/tap/



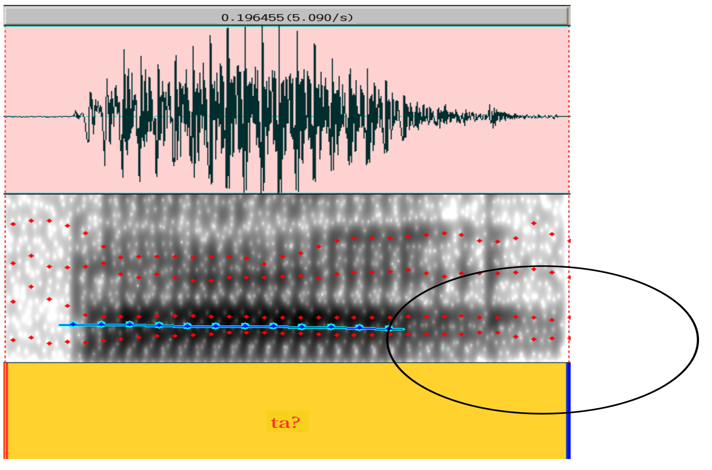
/tat/



/tak/



/taʔ/



allow /tap/, /kap/, /tat/, /tat/, and /kaʔ/, although it incorporates /tak/, /kak/, and /ta ʔ/. The hypothesis is that, as the Shantou dialect allows both /tak/ and /kak/, subjects would be able to detect /k/ more correctly than others, and /p/ and /t/ would be the least easy to identify. However, the results do not seem to support this. The lexicon does not constrain subject identification, at least in the case of /p/. However, the fact that the percentage of the correct identifications of /k/ and /ʔ/ are higher than those of /t/ seems to follow the prediction. Thus a further intrinsic analysis is needed.

Subjects may confuse /t, k, ʔ/ because the second formant transition is close to one another. However, if subjects are not confident about an answer, then the categorical perceptual theory plays a role, which means that the familiarity of the sound could be relevant, as both /ak/ and /aʔ/ are in the Shantou lexicon. This only explains why /k/ and /ʔ/ are more recognizable than /t/. On the other hand, even though /ap/ and /at/ are not existing rimes in Shantou, the subjects were able to detect /p/ at a correct rate that was better than that of /k/ or /ʔ/. The explanation lies in the fact that the in-variety of second formant transition precedes the lexical knowledge pool when listeners try to identify a sound. Once the nature of the acoustic information is adequate for subjects to categorize a sound, /p/ was successfully detected. Later on, they found that they could not categorize the rest, and lexical familiarity was pooled to help with identification. A similar claim was made by Tsukada (2006) in a study on native Australian English (AE) speakers and Thai—English bilingual speakers (TE). Because participant performance differed in discrimination tasks, the conclusion was that both first language (L1) transfer and phonetic sound realization ac-

Table 5 Existing Combinations Observed in Shantou and Tainan Dialects

	coda (Shantou)			
onset (row)	p	t	k	ʔ
t			x	x
k			x	

	coda (Tainan)			
onset (row)	p	t	k	ʔ
t	x	x	x	x
k	x	x	x	x

counted for patterns that elicited listener response. Our results not only confirm Tsukada's proposal, but go further to argue that acoustic character is first taken into account, and this is followed by sound discrimination based on first language knowledge.

We asked subjects in our study to detect which consonant they heard. This process reflects what the subject hears, but not necessarily the intrinsic nature of the signal. It has been reported (Werker and Tees 1984:59) that infants can discriminate all sounds before one year of age, but afterward only discriminate the sounds of their own language. It can be interpreted as the fact that the acquisition of one language is the process of losing the sensitivity to detect non-native sounds. Thus, subjects were 'deaf' to phonemes that do not exist in the coda position in their native language, of which /t/ is one example. Additionally, Ohala (1990:265) argued that, when making the recognition, people rely on the [release] feature for /t/ more than /p/ or /k/. This makes the unreleased final stop /t/ difficult to recognize. Jun (1995:10) also claimed that coronal gesture is articulated rapidly, resulting in a shorter transition that has the least influence on formants. In brief, the lack of adequate clear acoustic information and the absence of terminal contrastive systems are the two reasons that make /t/ less recognizable.

There were at least two parts where subjects were asked to identify the correct final stops. In the first, acoustic phonetics played a role in terminal plosive categorization. Formant transition could help subjects exclude the possibility of labial stops. The second was the phonological constraint. The ability of subjects to identify codas was constrained by their native language. If a subject's lexicon does not allow the combination of a vowel and a coda, the final stop cannot be successfully recognized. Of particular note is that, even though the native language of the participants may not allow a particular combination (e.g., /ap/), they would still be able to recognize /-p/, which was detected already in the first phonics level.

5. Conclusion and Further Research

The acoustic information of /-p/, characterized by F2, is the most important cue for the identification of codas because it differs significantly from the other three consonants /t, k, ʔ/. Shantou natives were unable to identify /t/, as it is a non-phonological sound in coda, while the identification of /k/ was as good as /ʔ/ in the second position in terms of percentage correct. In an earlier study (Ohala 1995), the shape of the Hindi VC formant transitions of stops was found to create many transitions that were similar when adjacent to specific vowels. Related information was listed as follows:

(1) Labials and dentals have similar transitions after /i/.

(2) Labials and velars are similar after /u/.

(3) Dentals and velars are similar after /a/.

This study confirmed the existence of a similarity between dentals and velars that follow /a/. Different vowels, such as /i/ and /u/, may reveal different results.

APPENDIX A

taʔ_1	kaʔ_1
taʔ_2	kaʔ_2
taʔ_3	kaʔ_3
taʔ_4	kaʔ_4
taʔ_5	kaʔ_5
taʔ_6	kaʔ_6
tap_1	kap_1
tap_2	kap_2
tap_3	kap_3
tap_4	kap_4
tap_5	kap_5
tap_6	kap_6
tat_1	kat_1
tat_2	kat_2
tat_3	kat_3
tat_4	kat_4
tat_5	kat_5
tat_6	kat_6
tak_1	kak_1
tak_2	kak_2
tak_3	kak_3
tak_4	kak_4
tak_5	kak_5
tak_6	kak_6

REFERENCES

Agresti, Alan. *An Introduction to Categorical Data Analysis.* New Jersey: John Wiley & Sons, Inc., 1995.

Blumstein, S. E. and K. N. Stevens. "Acoustic Invariance in Speech Production: Evidence from Measurements of the Spectral Characteristics of Stop Consonants," *The Journal of the Acoustical Society of America*, 66.4, 1979, pp. 1001-1017.

Chen, Matthew. "Cross-dialectal Comparison: A Case Study and Some Theoretical Considerations," *Journal of Chinese Linguistics*, 1.1, 1973, pp. 38-63.

Ciocca, V., L. Wong and L. K. H. So. "An Acoustic Analysis of Unreleased Stop Consonants in Word-final Position," *ICSLP*, 94, 1994, pp. 1131-1134.

Cohen, Ayala. "Dummy Variables in Stepwise Regression," *The American Statistician*, 45.3, 1991, pp. 226-228.

Harris, K. S., H. S. Hoffman, A. M. Liberman, P. C. Delattre and F. S. Cooper. "Effect of Third-Formant Transitions on the Perception of the Voiced Stop Consonants," *The Journal of the Acoustical Society of America*, 30, 1958, pp. 122-126.

Hay, J., J. Pierrehumbert and M. Beckman. "Speech Perception, Well-formedness and the Statistics of the Lexicon," in Local, J., R. Ogden and R. Temple eds., *Phonetic Interpretation: Papers in Laboratory Phonology VI.* Cambridge: Cambridge University Press, 2004, pp. 58-74.

Hosmer, D. W. and S. Lemeshow. *Applied Logistic Regression.* New Jersey: John Wiley & Sons, Inc., 1989.

Jun, Jongho. "Place Assimilation as the Result of Conflicting Perceptual and Articulatory Constraints," *West Coast Conference on Formal Linguistics* (WCCFL), 14, 1995. pp. 221-237.

Karlgren, B. *Zhongshangku hanyu yinyunkangyao* (*An Outline on the Phonology of Middle and Upper Old Chinese*). Jinan: Qilu Press, 1987.

Kawasaki, H. and J. J. Ohala. "Acoustical Basis for Universal Constraints on Sound Sequences," *The Journal of the Acoustical Society of America*, 68.1, 1980, pp. S33-S33.

Klatt, Dennis. H. "Software for a Cascade/Parallel Formant Synthesizer," *The Journal of the Acoustical Society of America*, 67, 1980, pp. 971-995.

Ladefoged, Peter. *A Course in Phonetics.* Orlando: Harcourt Brace, 1975.

———. *Vowels and Consonants: An Introduction to the Sounds of Language.* Massachusetts: Blackwell Publishing Ltd., 2003.

Lee, Sook-Hyang. "The Duration and Perception of English Epenthetic and Underlying Stops," *The Journal of the Acoustical Society of America*, 89.4, 1991, pp. 1999-1999.

Liberman, A. and I. Mattingly. "The Motor Theory of Speech Perception Revised," *Cognition*, 21, 1985, pp. 1-36.

Lieberman, P. and S. E. Blumstein. *Speech Physiology, Speech Perception, and Acoustic Phonetics.* New York: Cambridge University Press, 1988.

Lindblom, B. and O. Engstrand. "In What Sense is Speech Quantal?" *Journal of Phonetics*, 17, 1989, pp. 107-121.

Lin, L. L. and X. F. Chen. *Guangdong minfangyan yunyi yanjio* (*A phonetic analysis of the Guangdong Min Dialect*). Shantou: Shantou University Press, 1996.

Malécot, André. "The Role of Releases in the Identification of Released Final Stops," *Language*, 34.3, 1958, pp. 370-380.

Mielke, Jeff. "Interplay between Perceptual Salience and Contrast: /h/ Perceptibility in Turkish, Arabic, English, and French," in Avery, P. B. E. Dresher and K. Rice eds., *Contrast in Phonology Theory, Perception, Acquisition*. Berlin, New York: Mouton de Gruyter, 2008, pp. 173-192.

Ohala, John J. "The Listener as a Source of Sound Change," in Masek, C. S., R. A. Hendrik and M. F. Miller eds., *the Parasession on Language and Behavior*. Chicago: Chicago Linguistics Society, 1981, pp. 178-203.

———. "The phonetics and Phonology of Aspects of Assimilation," in Kingston, J. and M. Beckman eds., *Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press, 1990, pp. 258-275.

———. "The Segment Primitive or Derived ?" in Ladd, Docherty ed., *Laboratory Phonology II: Gesture Segmental, Prosody*. Cambridge: Cambridge University Press, 1992, pp. 166-183.

Ohala, J. and D. Feder. "Listeners' Normalization of Vowel Quality is Influenced by 'Restored' Consonantal Context," *Phonetica*, 51, 1994, pp. 111-118.

Ohala, Manjari. "Acoustic Study of VC Transitions for Hindi Stops," *Proceedings of the XIIIth International Congress of Phonetic Sciences*, 4, 1995, pp. 22-25.

Olive, J. P., A. Greenwood and J. Coleman. *Acoustics of American English Speech: a Dynamic Approach*. New York: Springer-Verlag, 1993.

Redford, M. A. and R. L. Diehl. "The Relative Perceptual Distinctiveness of Initial and Final Consonants in CVC Syllables," *The Journal of the Acoustical Society of America*, 106, 1999, pp. 1555-1565

Stevens, Kenneth N. "On the Quantal Nature of Speech," *Journal of Phonetics*, 17, 1989, pp. 3-45.

Stevens, K. N. and S. E. Blumstein. "Invariant Cues for Place of Articulation in Stop Consonant," *The Journal of the Acoustical Society of America*, 64, 1978, pp. 1358-1368.

———. "The Search for Invariant Acoustic Correlates of Phonetic Features," in Miller, Eimas ed., *Perspective on the Study of Speech*. Hillsdale: Erlbaum, 1981, pp. 1-38.

Tartter, V. C., D. Kat, A. G. Samuel and B. H. Repp. "Perception of Intervocalic Stop Consonants: The contributions of closure duration and formant transitions," *The Journal of the Acoustical Society of America*, 74.3, 1983, pp. 715-725.

Tsukada, Kimiko. "Cross-language Perception of Word-final Stops in Thai and English," *Bilingualism: Language and Cognition*, 9.3, 2006, pp. 309-318.

Tsukada, K., T. T. A. Nguyen and R. Roengpitya. "Cross-language Perception of Word-final Stops: Comparison of Cantonese, Japanese, Korean and Vietnamese Listeners," *International Congresses of Phonetic Sciences (ICPhS) XVI*, 2007. pp. 1781-1784.

Van Wieringen, Astrid. *Perceiving Dynamic Speech-like Sounds: Psychol-acoustics and Speech*

Perception. Doctor dissertation. Amsterdam: University of Amsterdam, 1995.

Wang, William S-Y. "Transition and Release as Perceptual Cues for Final Plosives," *Journal of Speech and Hearing Research*, 2.1, 1959, pp. 66-73.

Warner, N. and A. Weber. "Perception of Epenthetic Stops," *Journal of Phonetics*, 29, 2001, pp. 53-87.

Werker, J. and R. Tees. "Cross-language Speech Perception: Evidence for Perceptual Organization during the First Year of Life," *Infant Behavior and Development*, 7, 1984, pp. 49-63.

尾音的指認受制於聲學語音跟音韻限制

朱曼妮
清華大學語言所博士生暨德霖技術學院講師

林惠文
台北醫學大學生物統計暨研究資訊中心

摘　　要

一連串的研究顯示，分辨子音會基於其聲學的特徵，本實驗試圖發現受測者在只給一小段片段的母音過渡，其指認尾音的正確率受何種因素影響。十位汕頭的受測者，需要在 /p,t,k,ʔ/ 四個選項中，選出正確的答案，而汕頭是屬於不允許 /ap/ 跟 /at/ 的語言，結果顯示其正確率爲：/p/＞/k/, /ʔ/＞/t/。/p/ 在 /a/ 之後的正確率是最高的。我們的解釋是因爲 /t/，/k/ 還有 /ʔ/ 的第二共振峰在 /a/ 之後比較類似，都是上升的方向，而 /p/ 在 /a/ 之後的第二共振峰爲下降的，在受測者每次都只聽到每個母音過渡的一小片段的實驗設計之下，/p/ 的正確率顯著高於 /t,k,ʔ/。本研究結果證明第二共振峰對於受測者在指認母音上，有其不變性；而 /t,k,ʔ/ 之區辨在於其對於 /ak/ 跟 /a ʔ/ 屬於汕頭字彙之一環。

關鍵詞：汕頭，尾音，正確率，第二共振峰

(收稿日期：2008.11.13；修正稿日期：2009.2.6；通過刊登日期：2009.10.16)