HONG KONG SPOKEN CANTONESE DATABASE

Lydia K.H. So and Robin Thelwall

University of Hong Kong & Hong Kong Polytechnic

ABSTRACT

As the initial stage of collecting a structured sample of spoken Cantonese for a database to be used for speech therapy and pronunciation teaching purposes, the present paper discusses the design of a linguistic questionnaire and illustrates some early acoustic analyses of vowels based on the current wordlist.

INTRODUCTION

Cantonese, the standard variety among the Yue dialects of Chinese (Cheung K-H, 1986:9), is spoken by a total of over 40 million people in Hong Kong, Guangzhou and Macao (and overseas Chinese). Hong Kong Cantonese, for historical and political reasons, constitutes a distinct geolinguistic entity. Cantonese in general and Hong Kong Cantonese in particular, has received only limited phonetic investigation and published description. Phonological descriptions are more frequent and Cheung K-H.(1986) presents an admirable summary of preceding work, both phonological and phonetic. This work provides a masterly up-to-date interpretation of the phonology, proposing resolutions to a number of phonological/phonetic issues that arise from earlier descriptions, but inevitably provides limited phonetic data (since that is not its primary purpose). The present work is based on two sources of questions, first the data from speech therapy clinics in Hong Kong on developmental phonetic problems (So & Dodd, 1992), and second from a general comparative phonetic perspective.

The work of So & Dodd (1992) gives us a list of features such as: stopping, fronting, deaspiration, aspiration, /h/ deletion, final consonant or glide deletion, stop weakening, initial consonant deletion, gliding, vowel substitution and backing. (These terms are explained in So & Dodd.) The general phonetic perspective provides questions about such features as functional and non-functional nasality, the raising of vowels in the context of specific consonants, the lowering of vowels in the context of nasals, coronal articulations, anticipatory and perseverative coarticulation, vowel duration and syllable context, segment duration, syllable duration and the mora and the contribution these factors play in the rhythm structure of Cantonese.

Previous instrumental work on Cantonese

From an initial survey of the literature, instrumental work concentrates on VOT (Lisker & Abramson, 1964: one speaker); pitch analysis and synthesis (Fok, 1974 - ten speakers; Hashimoto, 1972 - one speaker; Vance, 1976 - five speakers; Vance, 1977 - synthesised tones based on values in Hashimoto, 1972; Vance & Walker, 1976 same data as Vance 1976; Ching, 1981 - eight speakers); and vowel formant data (Ching, 1981; Clumeck, Barton, Macken & Huntington, 1981 five children and eight adult speakers; Lee, 1985 - six speakers, three from Hong Kong, three from Guangzhou; Zee, 1991). While these studies are important are important as a base for further work, they are grouped around a relatively small range of issues and the total numbers of speakers analysed for any specific purpose is very limited. Even though the topic of VOT has been the focus of widespread investigation across the languages of the world since Lisker & Abramson's seminal paper, no major investigation of Cantonese VOT appears to have been carried out, and yet it appears clearly from the speech therapy evidence that problems in realising the aspirated/unaspirated opposition in Cantonese are an issue for which data from normal mature speakers would be of value. In Cheung (1986) there is a very interesting chapter on the interpretation of syllable structure and duration features in terms of a Mora. This is not supported by instrumental phonetic data, and it would seem of interest to carry out some basic measurement of segments/syllables in order to test this important hypothesis.

THE NEED FOR A DATABASE

Our primary need for data on Cantonese is to answer questions raised in speech therapy practice, and the teaching of English pronunciation which presupposes a reliable description of the phonetics of the learner's L1. The collection of data for these purposes alone would still allow other speech scientists to make use of them for other purposes, for example dialectology, sociolinguistic and perhaps aspects of NLP. While these other interests are not our primary reasons, we will try to take them into account in the collection of data and in terms of making the collected data accessible.

The linguistic tools

The linguistic tools are obviously of crucial importance, both to be able to answer questions of the sort outlined above, and also to allow other speech scientists to ask and answer questions from other points of view. A current dataset (Spoken Cantonese Operational Word List - SCOWL) follows on from the Cantonese Segmental Phonology Test developed by So (1991) for use with children attending speech therapy clinics. SCOWL has been considerably enlarged and can be read by the subjects, but at present is limited to including all segmental combinations within syllable/lexical units, it currently contains over 750 items. This will be supplemented by tonal oppositions and utterance length units for coarticulation and pitch interaction beyond lexical units. In addition a sample reading text and a 'spontaneous' text will be added (probably derived from a picture description task), so that some uniform connected speech data can be sampled.

The ultimate aim of the project is to collect a standard dataset of Cantonese (words in context, sentences, read texts and spontaneous speech) with simultaneous recording of a number of different phonetic parameters (audio signal, dynamic palatography, airflow and larynx signal) with perhaps some more narrow samples of fibroptic and x-ray data. Before we can be sure that the small sample of speakers is representative, we are planning to carry out more extensive audio recordings both for the full dataset and for specific subsets to establish the details of 'norm' and variation'. From the existing literature there are a number of such features that can already be identified for more wide-ranging investigation, but the audio samples to be collected will surely throw up further features. Current features known to be variables include: deletion of initial ŋ-, loss of labial-velarisation of initial velar stops, incidence of initial l/n alternation and the split of Tone 1 (High Fall) into High Fall and High Even (all discussed in Cheung K-H., 1986). Coarticulation on the syllabic level will be a focus of investigation as well as the interaction of lexical tone with utterance pitch and lexical tone with segments/syllables.

Instrumental Methods

The resources of the Department of Speech and Hearing Science, University of Hong Kong are available, supplemented by certain items of equipment purchased under the project funding.
Currently these comprise DAT recorders, Visipitch, LSI Speech Work Station, Laryngograph, Dynamic Palatography (Hardcastle/Millgrant Wells). Access to a NEXT workstation is available and IBM 486 and Vax systems. The Kay CLS and the full version of the Kay Spectrograph are on order as well as two Macintosh workstations.

THE TARGET SPEAKERS

In order to solve the problem of a suitable target group for investigation, we have chosen the 18-25 year-old group which we suggest is already stable in competence and performance and within which variation will merely reflect the normal range of idiolectal and sociolectal features which all linguistic description entails. We further assume that this group offers a plausible target for expected acquisition. A requirement of the target group is that they should have been born, raised and educated in Hong Kong with no significant period spent outside Hong Kong and in addition both parents should have been born and lived in Hong Kong for most of their lives. We can take selected speakers from

among the tertiary student population (though a sampling bias must be guarded against here in terms of educational/social class stratification).

How many speakers?

The final aim of our data collection is to use a range of instrumental techniques to record simultaneously on a small sample of speakers (at present proposed as eight). However, before we get to that stage we need to know something about norm and variation in the target age-group. We have started with one subject for detailed investigation to serve as an idiolectal base-line both for comparison with earlier studies and to provide a testbed for the development of the wordlists and sentences, and the correct functioning of the instrumental methods.

The initial subject has recorded the current version of the wordlist and some acoustic analyses of this data are presented below. At the same time a second research assistant has recorded the wordlist and will shortly assist in collecting the same dataset from a proposed group of approximately 64 speakers, half male, half female, covering students both at Hong Kong University and Hong Kong Polytechnic. Since the Polytechnic has students taking a range of vocational courses it is expected that these students will extend the social range of speaker. After this stage, the selective analyses of the audio recordings from this larger sample, plus the detailed instrumental work with one or two individuals, should allow us enough understanding of the Cantonese of this age-group to set the criteria for the selection of the final eight for the instrumental study. It is to be expected that a number of additional investigations for very selective features will be carried out with other informants.

CURRENT STATE OF THE PROJECT

The project is funded for two years (Hong Kong UPGC Grant No. HKU 242/92H) and a part-time research assistant started in August 1992. This RA will be the subject for the initial application of the instrumental techniques. His data is represented in the formant chart below. A second part-time RA is being appointed to start the audio survey. He will also be trained to carry out some of the transcription work. Acoustic analysis of the first subject (RW, male, aged 20) is given in table 1. This is designed to show that even at this early stage we can be analysing limited data and comparing them with existing descriptions in order both to assess their 'normality' and to formulate further questions.

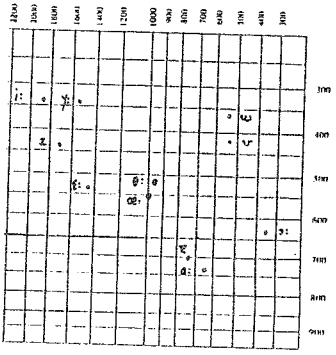
Tokens were selected from complete set; excluded tokens were those where the syllable ended in a glide; Tone was not controlled.

		F1	F2	F3	
Vowel [i]	Average =	336.62	2,193.45	2,990.55	29 tokens
Vowel [y]	Average =	348.54	1,923.15	2,487.54	13 tokens
Vowel [e]	Average =	442.53	2,136.60	2,705.64	15 tokens
Vowel [ɛ]	Average =	540.79	1,975.89	2,686.94	19 tokens
Vowel [ɵ]	Average =	526.33	1,503.50	2,609.67	6 tokens
Vowel [œ]	Average =	551.58	1,551.67	2,519.50	12 tokens
Vowel [a]	Average =	731.64	1,389.56	2,557.79	39 tokens
Vowel [o]	Average =	629.54	997.17	2,387.71	24 tokens
Vowel [ʌ]	Average =	715.57	1,438.17	2,530.24	35 tokens
Vowel [U]	Average =	419.71	955.50	2,213.93	14 tokens
Vowel [u]	Average =	365.25	902.88	2,677.25	8 tokens

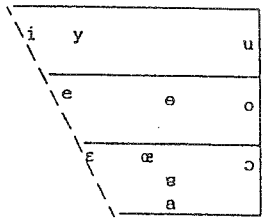
Table 1. Average formant values for steady state of one subject (RW)

These values are presented on a vowel chart below, together with vowel charts from Cheung (1986)

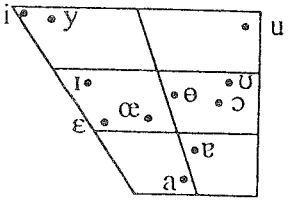
and Zee (1991). While neither chart claims acoustic support, though Zee has carried out acoustic analysis of Cantonese, there are some initial questions that we can raise, even on the basis of one speaker. The relative location of all the back vowels is different in all three charts. This provides one topic for further investigation in our survey. The relative agreement between our chart and Zee's for the central rounded vowels appears to go against Cheung's. One question is are there any phonological patterns that support Cheung's proposed patterns? Another area is that of vowel length. In Cheung's work only [ɪ], [ʊ] and [ɐ] are classified as short, though there is some discussion of shortening of long vowels in some contexts. This can be easily measured with the present data.



So & Thelwall (1992)



Cheung (1986)



Zee (1991)

Figure 1. Cantonese Vowel Charts

REFERENCES

Cheung, K-H. (1986) *The phonology of present day Cantonese*, unpublished PhD dissertation, (University College: London).

Ching, Y-C.T. (1981) *Communication of lexical tone patterns in Cantonese*, unpublished PhD dissertation, (University College: London).

Clumeck, H., Barton, D., Macken, M.A. & Huntington, D.A. (1981) *The aspiration contrast in Cantonese word-initial stops: data from children and adults*, Journal of Chinese Linguistics, 9, 211-225.

Fok, C.Y-Y. (1974) *A Perceptual Study of Tones in Cantonese*, Centre for Asian Studies, University of Hong Kong: Hong Kong).

Hashimoto, O-K.Y. (1972) *Studies in Yue dialects. 1. Phonology of Cantonese*, (Cambridge University Press: Cambridge).

Lee, H-T.T. (1985) *The phonetic quality and long/short distinction of Cantonese vowels*, (in Chinese), Fangyan, 1, 28-38.

Lisker, L. & Abramson, A.S. (1964) *A cross-language study of voicing in initial stops: acoustic measurement*, Word, 20, 384-422.

So, L.K.H. (1992) *Cantonese Segmental Phonology Test*, Department of Speech & Hearing Sciences, University of Hong Kong, Hong Kong

So, L.K.H. & Dodd, B. (1992) *Phonologically Disordered Cantonese-Speaking Children*, University of Queensland, Brisbane)

Vance, T.A. (1976) *An Experimental Investigation of Tone and Intonation in Cantonese*, Phonetica, 33, 368-392.

Vance, T.A. (1977) *Tonal Distinctions in Cantonese*, Phonetica, 34, 93-107.

Vance T.A. & Walker, C.A. (1976) *Tone and Intonation in Cantonese*, Canadian Linguistic Society, 12, 640-655.

Zee, E. (1991) *Chinese (Hong Kong Cantonese)*, JIPA, 21.1. 46-48