

Hong Kong University of Science and Technology

HUMA 6000G
Spring Semester 2012

Topics in Chinese Phonetics:
Forensic Voice Comparison in Cantonese

廣東話法庭語音比較

[k^wɔŋ¹tuŋ¹wa:¹ fa:t¹th¹ŋ¹ ɥy¹ljəm¹ pɛi¹ kau¹

fə,ɪɛnzɪk 'vɔɪs k^həm,p^hæ.ɪsən ɪŋ k^hænt^hə'ni:z]

COURSE INTRODUCTION

In Forensic Voice Comparison (FVC) a recording of an unknown voice – usually of someone committing a crime – is compared with a known voice – usually of a suspect. The aim is, of course, to help in the decision of whether the suspect said the incriminating speech.

In this course you will learn how voices are compared forensically, and you will learn in the best possible way – by actually doing it. The way that voices are properly compared forensically differs a lot from the typical depictions in the films (i.e. movies) and TV. Forensic Voice Comparison has advanced enormously in the last decade, and some testing has been done using Standard Chinese voice samples. We will be investigating how well it can be done using spoken Cantonese, which has not been done before. Students therefore will have an opportunity to participate in a unique and exciting event!

In the course students will collect recordings from different speakers, extract and measure acoustic features from their speech, and conduct an experiment to discriminate between speech samples from the same speaker and speech samples from different speakers using those features. Different students will compare speech samples with different features, and towards the end of the course they will be showed how to combine the results from all the different features to get a better overall discrimination.

In order to do Forensic Voice Comparison properly, you must have a good understanding of aspects of several main areas of knowledge – most of them not easy – and in the course we will therefore focus on these. The first and most important question is one of *applied probability in Logic*: how do you evaluate the probability that a hypothesis is true, given the evidence adduced in its support? (In FVC: How do you correctly answer the trier-of-fact's question: what is the probability that two speech samples have come from the same person, given the degree of similarity, or otherwise, between them?) Knowing how to evaluate hypotheses from their evidence is of course of extremely wide application outside of forensics, and it is hoped that it will be of use to students not just in this course but in all their studies, and in life after university.

A second question is: how do you quantify the difference between the speech samples?

To decide this you first have to know about the subdisciplines of Linguistics dealing with speech: *Phonetics* (especially the acoustics of speech and how they are produced), and *Phonology* (how speech sounds are organized). A certain amount of *statistical* and *signal-processing* knowledge is also necessary, to understand how to obtain and process your acoustic data and to understand how the formulae are supposed to work.

It can be seen that FVC brings together knowledge from several different areas of both the humanities and science in order to accomplish an important goal, and thus it is an excellent subject choice for the Humanities Division of a University of Science and Technology.

A course website will be used which includes unique materials designed by the lecturer to develop students' analytical skills. The lecturer is an acknowledged expert on forensic voice comparison.

“LEARNING OUTCOMES”

1. You will learn how Forensic Voice Comparison is done properly.
2. You will learn about the logical evaluation of evidence, in particular how to reason under uncertainty.
3. You will learn about the nature of some Cantonese speech sounds: how they relate to the speaker producing them and how they can be measured acoustically.

LECTURER: Dr. Phil Rose.

LECTURES/ PRACTICAL/TUTORIAL: Three per week.

VENUE: 3416

TIME: Saturday 10.00 – 12.50

WORK

1. **Reading and understanding** Due to the multidisciplinary nature of the topic, and the fact that it hasn't been done before, there is no single appropriate text-book. Separate readings going into specific topics in greater detail, and of course further references, will be made available regularly during the course. Make sure you set aside enough time for reading, as the material is often not easy, and requires more than one reading.
2. **Attendance** Obviously, attendance at lectures is important, and this is doubly so for a course like this, which is strongly cumulative, and which cannot rely on a lot of pre-existing information to be rote-learned: it will be very difficult to understand a lecture if you have missed the previous one. Also I expect lots of questions to be raised and discussed by all of us during contact hours, and after, and the resulting discussion is always informative and worth-while from the point of learning. So, if

you are unable to make a lecture, please find out soon, and understand from your fellow students who were there what we talked about.

3. **Assessable work** There will be three assignments counting 20% each, and one final larger project counting 40% to your total mark. The first two assignments constitute necessary steps towards the third in the sense that you have to complete them successfully before you can attempt the third.
 - The first assignment will involve data collection, preparation, and description. Students will have to record about 10 speakers each so that we have adequate experimental data to test. It is extremely important to get good recordings (and you will be assessed in part on this). Instructions on how to do this properly will be given out early in the course, and there will be opportunity to practice (indeed it is essential to practice first). Students can use recordings of their own voice (if they are native Cantonese speakers). If you intend taking the course, *you should start looking around for and recruiting possible speakers right now!!* They need to be male and native Cantonese speakers (fathers, brothers, cousins, male friends, male friends of friends ... they are all good, but they will need to be able to come to the UST to be recorded (so perhaps fellow students is the best option). Each speaker will need to be recorded twice, with a separation of about two weeks. (What we understand by *Cantonese* is important – we don't want to make the experiment too easy by including, say, a Taishan 台山 speaker. On the other hand, someone from Canton 廣州 might be ok. For the moment, then, look for speakers of Hong Kong Cantonese, and ask me if you would like to include a speaker of a different Yue 粵 variety.)
 - The second assignment will involve measurement and statistical description of selected acoustic features from the recordings you have made, to show for example how much your different speakers differ and how much they are the same, in the particular feature you have chosen to use. Your measurements must be prepared so as to input into the final stage of the analysis in assignment three.
 - In the third assignment, Likelihood Ratios are estimated and the actual forensic voice comparison tested for the particular feature under consideration, and evaluated with the conventional metrics. You will see how well your feature can discriminate between speech samples from the same speaker and speech samples from different speakers.

In the larger project, you get a chance to demonstrate how well you can apply your expertise. Students will be given data and information from a real forensic voice comparison case and will write a report describing the full analysis, ending of course with your finding for the court, and an explanation of what it means. In the sense that it will be completed after the teaching term ends, the larger project counts as a quasi-exam. It is unlike an exam in that you will be able to ask for my advice and help.

SYLLABUS

We have to study the following topics.

1. **The basic ideas:** Forensic Speech Science, Forensic Voice Comparison, Conditional Probability, Evidence, Hypotheses, Bayes Theorem, Likelihood Ratio, Prior Odds. Reference Sample, Testing.
2. **Data acquisition:** Requirements for adequate testing material. Speakers, Corpus, Replicates, Elicitation, Recording, Editing.
3. **What to compare:** Semiotic model of voice, Acoustic Theory of Speech Production.
4. **Software:** If you can't already, you will need to learn how to use basic speech analysis software to extract acoustic features. As a default, we will use *Praat*. If you are used to using other software, that is fine, but check with me first. Download *Praat* from:

http://www.fon.hum.uva.nl/praat/download_win.html

- Another software useful for editing speech is *Audacity*, downloadable from:

<http://audacity.sourceforge.net/download/>

- It's a good idea to download these now and start playing with them (although you will of course be shown how to use them during the course). The most important thing to be able to do at first is read-in .wav format audio files.

You will also need to be able to manipulate data – get it in and out of the FVC analysis and processing software, and ideally you need to be able to write a little code to tailor things to your own analysis (this is the only way to get the computer to do what *you* want it to do, rather than the other way round). As a default we will use *R* (*Project for Statistical Computing*), as much of the processing code is written in it. If you are used to coding in Matlab, that is also fine, as LR code also exists in it. Download *R* from:

<http://www.r-project.org/>

(Find the “Getting started” window and click on download R.)

5. **Sampling speech acoustics:** Different types of features.
6. **Likelihood Ratio estimation:** Modelling distributions, Univariate LR, Multivariate LR, UBM-GMM LR.
7. **Evaluation of Performance:** Tippett plot, Log Likelihood Ratio Cost (*Clr*), Equal Error Rate (*EER*), Accuracy, Precision, Validity, Reliability, Calibration.
8. **Fusion:** Logistic regression.
9. **History:** The new paradigm of forensic evidence evaluation and its emergence.

Due to the multidisciplinary and practical nature of the course, these topics cannot all be approached in succession, and some will be interwoven (so that we might for example spend an early week discussing what a voice is, how to record speakers, and how to write some basic code in *R*).

Other matters

- **COLLABORATION, COPYING & PLAGIARISM**

Please be aware that **all your work must be completed entirely on your own, and not in collaboration with others. Collaboration, and any form of copying, entails severe penalties, including forfeiting the mark for the problem, and possible failure of the whole unit.** Please also note that, as in any scholarly work,

You must give credit for material you have obtained from other sources, by citing a reference, for example.

Failure to do this – presenting someone else's work/ideas as your own – is called *Plagiarism*. It attracts severe penalties: don't do it. If in doubt, ask! If you are taking the course, I will assume that you have read this.

OK, THAT'S IT. SEE YOU IN JANUARY.

[ʔar'ouɸɕjuŋ'dʒɔiðə'kʰɔ:s]