

FIGURE 1. View of the excel data sheet

## 1. Short study of Randall's Mandarin F0 contour data

Randall and Xin Chuan put together a nice piece of data for the study of Mandarin tone. The data file contains thus far 361 recordings and evaluation of pitch contours for most of the combinations of tones embedded in a phrase. Here are some first results extracted from these data.

### Description of the data

Each record contains the following information: (i) A three-digit string which represents a sequence of tones. (For example 232 means, second tone followed by 3rd tone, followed by second tone). (ii) A string in pinyin with tone information, representing the test phrase (independent document in Chinese characters is accompanied). (iii) A record of the measured F0 data over a range of the syllable subject to measurement. This record also contains the time since beginning of the utterance. (iv) A screen plot of the display of praat.

### Preprocessing

The numerical data and string information from the excel file were saved as ASCII text files with tabulators. A script was written in C that reads the text file and generates an output more suitable to machine processing, either with Matlab or otherwise.

Additional information was extracted and written to the file. For Matlab processing the structure of each record looks like this:

```
dat{1}:
  ltone: 1
  tone: 1
  rtone: 1
  dur: 0.1700
  start: 1.7837
  len: 18
  text: 'zhei4ge4shang1chang3you3hen3duo1xin1xian1shu1cai4.'
  lsyl: 'duo1'
  syl: 'xin1'
  rsyl: 'xian1'
  lask: 'duXn'
  ask: 'xJE'
  rask: 'xJA'
  F0: [1x18 double]
```

In this data structure, the tags have the following meaning:

ltone, tone and rtone represent the tone information of the syllable under study and the left and right syllable.

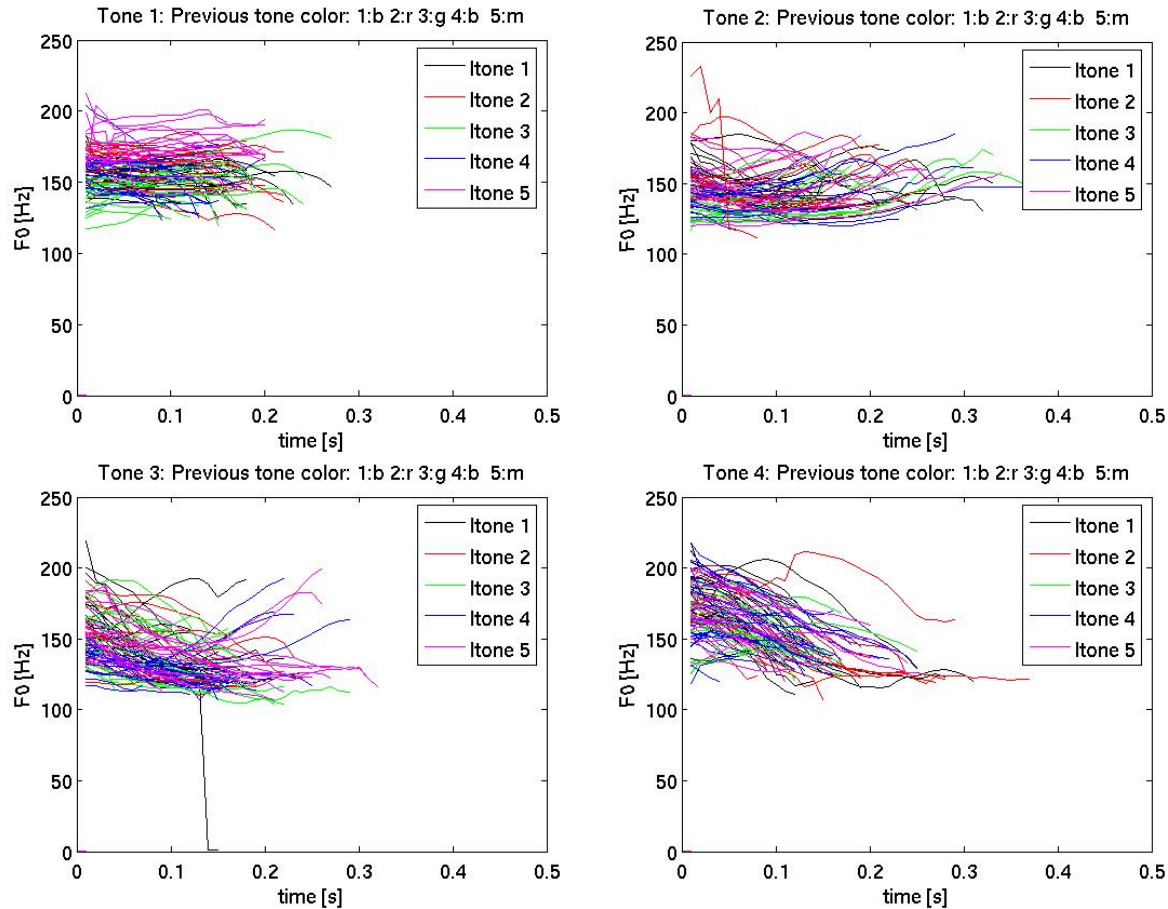


FIGURE 2. For the four tones the F0 contours are overlaid and labeled with 5 different colors depending on the preceding tone. See legend of the figure, the neutral tone is the 5th tone (magenta)

The *dur* parameter specifies the duration in seconds of the F0 record, and *start* is the point in time where the F0 record starts, relative to the beginning of the utterance.

The *len* parameter specifies the number of F0 values.

The *text* is the spoken text in pinyin with tone information.

The tag *syl* specifies the syllable under study with tone. *lsyl* and *rsyl* are the pinyin and tone information of the preceding and following syllable. *lask*, *ask*, and *rask* are the ASKY representation of the phonemes in the three syllables. Finally, *F0* is the record of the (time-equidistant) values measured by *praat*.

### Initial processing and display

Simple scattergrams often tell much of the story, so that's where I started. A simple program was written that allows to overlay all F0 contours. The first one is shown in Fig. 2.

There is of course an alignment problem: If the first phoneme of an utterance is voiceless, there are no F0 measurements, and the plot of the F0 contour should actually be delayed by the length of the unvoiced onset of the syllable. So in Fig. 3 a delay of 75 ms was assumed for all syllables which start with a voiceless consonant. This is of course a very rough correction.

### Considerations

These data give already some good hints at how to improve rules for tone synthesis. Including the proper position of the start of syllable might make the variance of the observed F0 contours smaller and it may be possible to see clearer patterns in the dependency of the preceding and following syllable. I also didn't really look into the dependency of syllable durations, which of course vary considerably.

Friday, March 10, 2006. Reiner

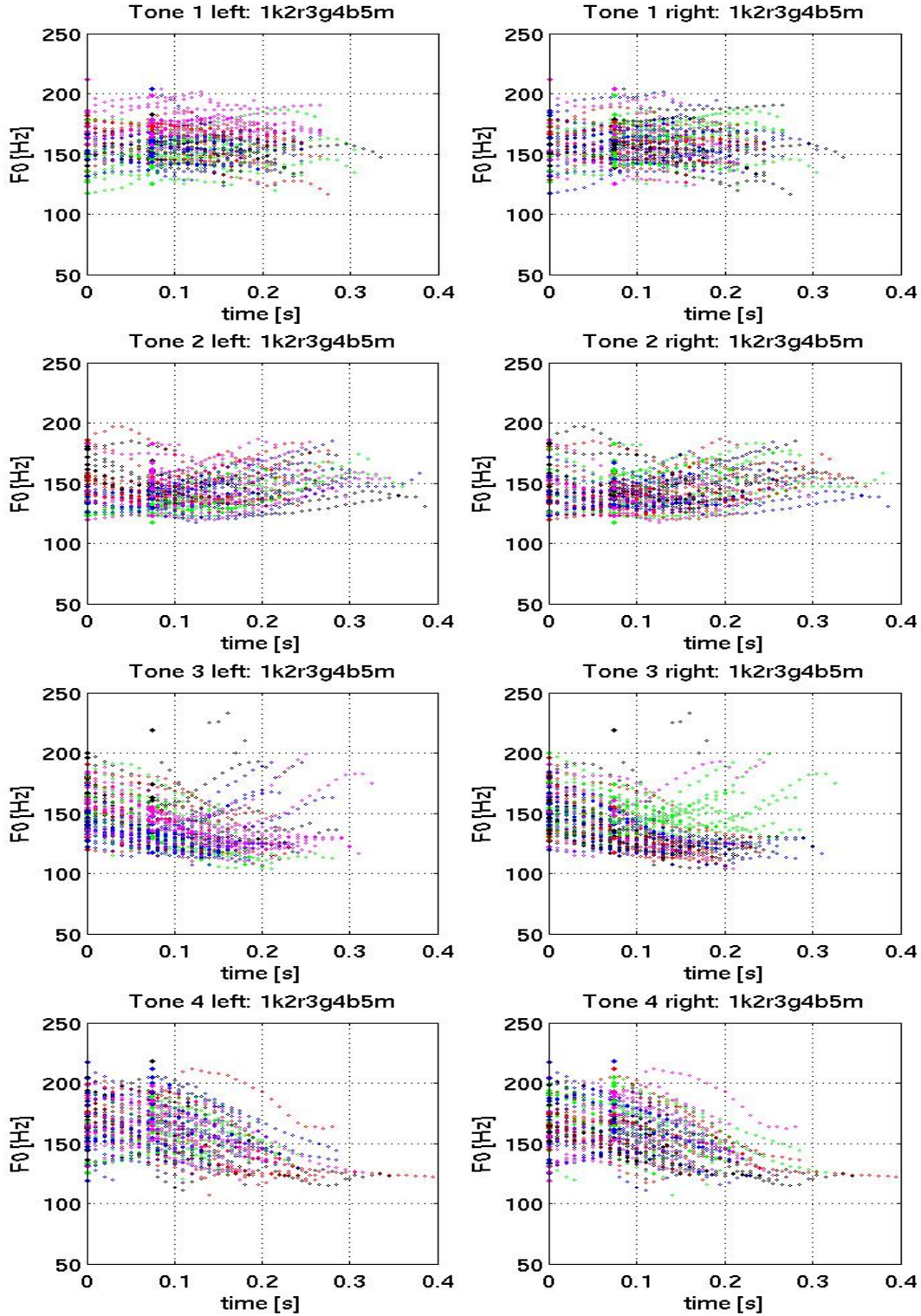


FIGURE 3. For the four tones the F0 contours are overlaid and labeled with 5 different colors depending on the preceding tone (left) or following tone (right). F0 contours of syllables which start with an unvoiced consonant are displayed with a delay of 75 ms. The colors for the left or right neighbor tones are: black for 1st tone, red for 2nd, green for 3rd, blue for 4th, and magenta for neutral tone.