F_0 Behaviour in Mandarin and French: An Instrumental Comparison

R. Sneppe and V. Wei
Brussels, Belgium

1. Introduction

Learning Mandarin Chinese is almost proverbially difficult for adults and this is not restricted to 'foreigners'. Native Chinese raised in a different dialect, e.g. Cantonese, have a problem in speaking Mandarin, see Tse (1974). What are the reasons for this difficulty? Three types of explanations have been advanced.

First Tse (1974) argues for the simple interference hypothesis. Cantonese have problems when learning Mandarin because two tone contours that are in free variation in Cantonese, have a lexical status in Mandarin. This is the main source of confusion.

A second hypothesis has been advanced by Chen (1974). Comparing the native speech of Chinese and Americans, he reaches the following conclusion: 'When the test subjects spoke their native language, (...) the average (pitch) range of the four Chinese subjects was 154% wider than that of the four English-speaking subjects... an English speaking person... should widen his normal pitch range at least 1.5 times if he wants to successfully learn to speak Chinese.

Finally Eady (1982), comparing American English and Mandarin arrives at a very different conclusion. '... for both the Mandarin and English speakers, the compass[1] of the voice (which is directly related to standard deviation) was a good indication of the range of F_0 values produced. The lack of significant difference between the two language groups for the variable SDFF (Stand. Dev. of Fund. Freq.) indicates that the F_0 values for the Mandarin speakers varied over a range that was approximately equal to that of the English.' Instead of a difference in pitch range, Eady records a statistically significant increase of F_0 fluctuation (as a function of time and of the number of syllables per second) for Mandarin speakers.

These quite different and sometimes conflicting hypotheses call for more investigation. Yet another comparison is presented here. Using French, a language known for its relatively 'flat' intonation, and Mandarin, stretches of continuous speech were analyzed with respect to F_0 behaviour.

Cross-language comparisons using different speakers are difficult to interpret. It is indeed a very intricate business to disentangle the various factors and their influences. Hence the usefulness of bilingual speakers. In this study a single bilingual French-Chinese speaker was used. If there are differences between Mandarin and French, these should be obvious in his speech.

2. Method

2.1. Subject

One of us being Chinese-French bilingual, the choice of the subject was obvious. The definition of bilingualism adopted here is: on the phone, no native speaker of Chinese or French would mistake the bilingual as a foreigner. Off the phone, any accent should be attributed to intra-dialectal variation rather than to foreign accent.

2.2. Recording procedure

Two texts were read. In French it was a fairy tale: 'La Reine des Neiges.' (The Queen of Snow). In Chinese it was a monologue taken from De Francis' Intermediate Course (Lesson 2). Both texts were read unemotionally and at normal pace. The recording was done under low level background noise.

2.3. Analysis procedure (see figure 1)

The filter is set at a given frequency. Any voiced segment coming out of it starts a time window (here 10 ms.). During this time, each period, transform-

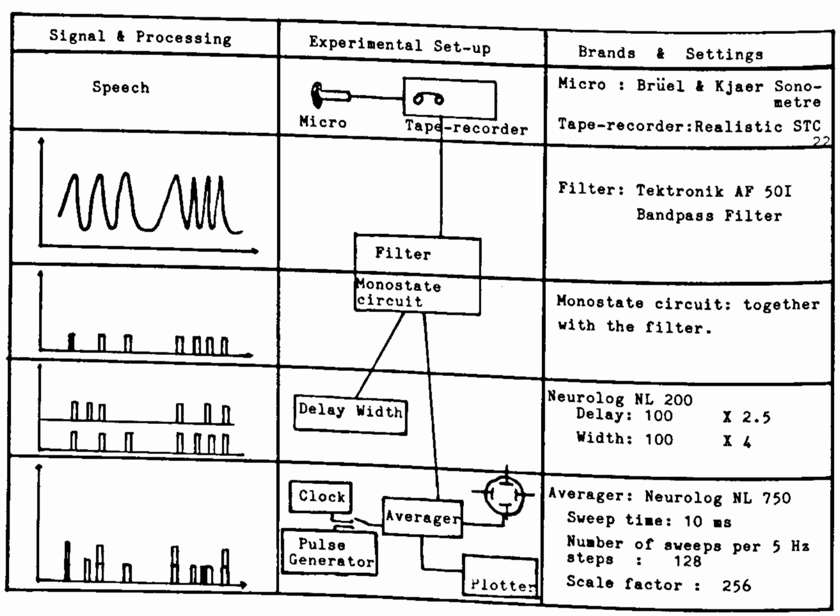

Figure 1. Description of the experimental set-up.

ed as a normalized pulse, is recorded in one of the 256 memory blocks of the averager. When the window has opened 128 times the averaging stops and the process is restarted, with the filter 5 Hz lower or higher depending on whether the filter is sliding towards 120 Hz or 200 Hz in 5 Hz steps. This gives two curves for each language. One called the increasing curve (i.e. when the filter is sweeping upwards to 200 Hz) and the other is, of course, the decreasing curve. Motivation for the upsweep downsweep is discussed in 4.2.

3. Results (see figure 2)

From the curves the following information can be deduced:

1. Weighted mean F_0 for Chinese upsweep: 173 Hz Standard Deviation: 29 Hz
   Weighted mean F_0 for Chinese downsweep: 173 Hz Standard Deviation: 31 Hz
   Weighted mean F_0 for French upsweep: 166 Hz Standard Deviation: 22 Hz
   Weighted mean F_0 for French downsweep: 166 Hz Standard Deviation: 21 Hz
2. Chi-square tests were run, giving the following results: the two Chinese curves were similar with a probability under .1; the two French curves

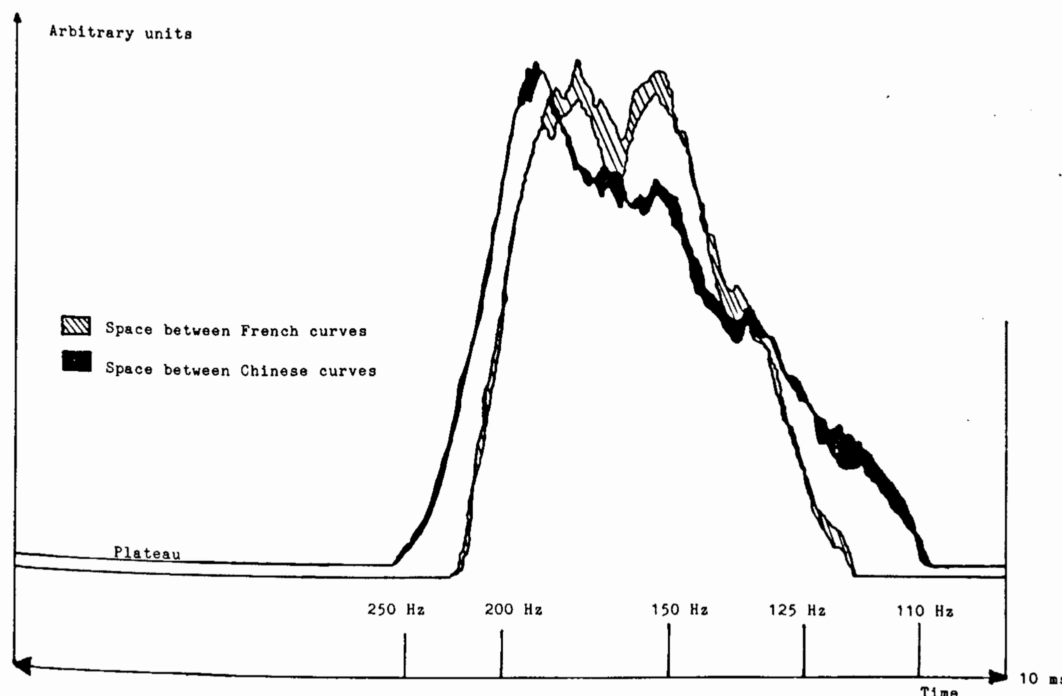

Figure 2. F_0 histogram, expressed in arbitrary units vs. frequency (time).

were also similar with a probability under .10; finally each Chinese curve was different from each of the two French curves with a probability over .99 . From these results, it is clear that Chinese F_0 behaviour is different from the French one.

4. Discussion

Various factors influence the curves presented here. They will be

1. Choice of the window width: 10 ms was chosen because the speaker did not have a significant contribution of pitch under 110 Hz.
2. Filter: It was found that changing the filter value resulted in substantial curve modification. In order to reach a curve closer to reality, it was decided to resort to a 5 Hz upsweep and downsweep from 120 Hz to 200 Hz for each language. This coupled with the same number of averaging gave a more balanced picture.
3. Segmentation: The procedure adopted here cuts up the speech in groups of 128 windows of 10 ms. This segmentation is not the same for the upsweep and the downsweep because first we did not make sure to start at exactly the place on the tape; secondly, frequency repartition in the spoken stretches are not identical, hence for some frequencies (especially the ones close to the mean) the 128 windows are recorded in less time than for other frequencies. This means that the upsweep and downsweep do not yield the same segmentation. It is as if two different texts were read, one for the upsweep and another for the downsweep. This fact diminishes the influence of the particular texts.
4. Pulse width: this was chosen here to give a reasonably detailed picture.
5. Plateau: Pulses are created by a voltage, however the absence of pulse does not correspond to an absence of voltage; even in the absence of pulse there is a slight continuous voltage. This explains the presence of the plateau. The plateau height is proportional to the number of averaging done.
6. Number of averaging was set at 128 pulse trains. (i.e. the window opened 128 times). We found this to be a convenient value.

The factors above mentioned should be put in the context of this important argument: the subject's reading of French and Chinese texts was submitted to the same treatment. If any difference can be detected, it has to be attributed to language difference.

5. Conclusion

From the results obtained, it is clear that French and Chinese have different F_0 behaviour. Pitch range is greater in Chinese than in French. It can be tentatively inferred that French speakers ought to increase their pitch range by 1/3.

The procedure used here does not allow inferring any information as to the rate of fluctuation on Eady's sense (Eady 1982). This will be a topic for further investigation.

[1] voice compass: the stretch of frequencies to standard deviations on either side of the mean F_0.

References

Chen, Gwang-tsai (1974). The pitch range of English and Chinese speakers. *J. of Chinese Linguistics* **2** 2, pp. 159-171.

De Francis (1967). *Intermediate Chinese.* Yale University Press.

Eady, S.J. (1982). Differences in the F_0 Patterns of Speech: Tone Language vs Stress Language. *Language and Speech,* **25**, 1, pp. 29-42.

Tse, J.K.P. (1974). An Instrumental Study on the Difficulties with Respect to Tones encountered by Native Speakers of Cantonese when learning Mandarin Chinese. *Education and Culture Monthly*, pp. 62-66, Taipei.