Timing Expectancies in Japanese Speech
Abstract - 2nd International Workshop on Language and Speech Science

Michael C. Brady†& Nick Campbell‡
†Indiana University, USA, ‡NiCT/ATR, Kyoto Japan
mcbrady@indiana.edu, nick@atr.jp

1 Introduction

Gaining a command of the timing flow of a language is essential to being able to speak and understand what is spoken. It has been claimed that a 'proactive memory' for when the transitions into and out of vowels will occur in a speech signal can help us to attend to and parse the information in a speech signal [1, 2].

This paper describes and evaluates a method for analysing temporal structure in Japanese speech based on vowel onset patterns. In this method, vowel onset locations are first detected in a speech signal, and then a spectral timing model is used to evaluate periodicity in the event sequences.

The spectral model is built of a bank of resonators in where each resonates with a distinct natural period in response to the input signal [3]. Output from the model is the sum of responses from the resonators. When local periodic structure exists in the input signal, the resonators with a related natural frequency become more active. If the signal continues to have periodic structure, newly arriving events are better anticipated and the corresponding resonators are reinforced and remain active. However, if an event occurs at a time when active resonators are out of phase, those resonators will absorb the input rather than respond to it, and be correspondingly weakened. The spectral model boosts periodic events and filters non-periodic events.

2 Data & Analysis

Manually annotated speech recordings from the talking-book "Usotsugi Tamago" read by Ms T. Kuroyanagi, a well-known personality and broadcaster were analysed according to the above model. Figure 1 illustrates output from the program showing how the vowel onsets in the word "tamago" have been detected as more periodically salient.

The full paper contains an analysis of periodicity in vowel onset positions, concentrating on the word "tamago" of the title, showing the extent to which they appear in more or less periodic contexts.

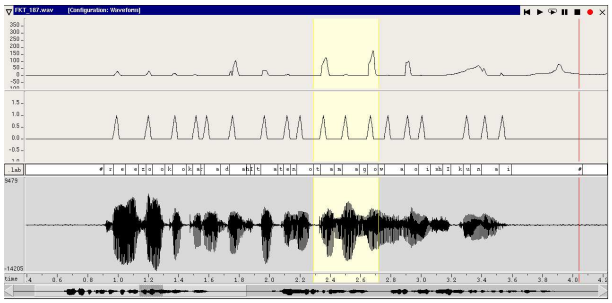

Figure 1. *Plots of the output, showing periodicity estimates for each vowel, vowel onset positions, and manually-assigned phone labels, aligned with the speech waveform*

3. Conclusion

By passing a pulse-train of voiced vowel onset positions through the resonator bank, we can observe resulting 'periodicity scores' as exemplified in the figure above. Output from the model will indicate statistical regularities in the speech, and the degree of activation measured for each vowel onset can be used as an indication of its rhythmic salience.

The paper presents an analysis of data from a large number of sentences in broadcast read speech, showing the extent to which the reader has placed key words in more periodic timing slots.

References

[1] Large, E. W., Jones, M. R. 1999. "The dynamics of attending: How we track time varying events". Psychological Review 106, 119159

[2] Brady, Michael C., and Port, Robert (2007) "Quantifying vowel onset periodicity in Japanese" Proc. 16th ICPhS Saarbrcken

[3] Brady, M. C., (2006). "Adaptive resonance situated for articulatory speech learning and synthesis." Proceedings of the 9th International Conference on Development and Learning.