Dynamic voice source parameters in a female voice

Karlsson, I.

journal: STL-QPSR
volume: 30
number: 1
year: 1989
pages: 075-077

http://www.speech.kth.se/qpsr
DYNAMIC VOICE SOURCE PARAMETERS IN A FEMALE VOICE

Inger Karlsson
Dept. of Speech Communication and Music Acoustics, KTH

Abstract

Dynamic variations of the voice source in speech have been studied for a female speaker. The speech material consisted of syllables and sentences with varying stress patterns. The speech signal was inverse filtered and then a parameter model of the voice source was fitted as closely as possible to the inverse filtered glottal pulses. The four parameter glottal source model, the LF-model, was used. This model is described by Fant, Liljencrans, & Lin, 1985, and is shown in Fig. 1.

INTRODUCTION

In the work on producing more natural synthetic speech and especially giving the synthesizer different voices, the description of the voice source forms an important part. Accordingly, voice source data are being collected from different speaker types. In this paper some data on a normal female voice will be presented. As the so called LF-model, Fant & al., 1985, is implemented in our text-to-speech system, the parameters of this model are utilized in the description. The model is presented in Fig. 1.

Fig. 1. Glottal air flow, upper curve, and differentiated glottal air flow, lower curve. The parameters of the LF-model are indicated.

\[ P_0 = 1/T_0 \]
\[ F_g = 1/2t_p \]
\[ r_g = T_0/(2t_p-100) \]
\[ r_k = t_o/t_p-100 \]
\[ r_a = t_o/T_0-100 \]
\[ F_a = 1/2\pi a \]
open quotient = \((t_p+t_o)/T_0\)
SPEECH MATERIAL
A female speaker was recorded reading lists of sentences and syllables. From these readings the following material was chosen for analysis:

Sentences
[ɪnɪˈdɛtərɔːθənˈdrɛd] (= not in this century),
[ɪnəˈdɛləsəˌwɪtˈmɪrktaɪd] (= a perfectly splendid idea),
[ˈvɪltɪbəˈholə:ənɔm] (we want to keep him)
The last sentence was pronounced with three stress patterns, with main stress on the word [vɪlt], on the word [hələ] and on the word [ənɔm].
The same sentence material was used by Gobl, 1988, in his study of male voices.

Vowels
The nine long Swedish vowels were pronounced in isolation and preceded by [h], that is the vowels were pronounced with hard and soft voice onset.
The female speaker was judged by a speech therapist to have a normal, somewhat tight and sonorous voice, see Karlsson 1988. In that paper, the speaker was signified by F1.

![Graphs showing dynamic voice source parameters](image)

**Fig. 2.** Dynamic voice source parameters for the female speaker, left part and a male speaker, right part. The male data are from Gobl 1988. The female $F_a$ values are higher than for this male only in the stressed rendering of the word [bɛhələ], in other cases and for the other two male speakers in Gobl's study, $F_a$ is about equal.
Analysis
The speech material was inverse filtered glottal pulse by glottal pulse by an interactive computer program and for each inverse filtered pulse the LF-model was fitted.

RESULTS

Sentence material
The dynamic variations of the voice source parameters studied, namely $E_e$, $n_k$, $r_g$ and $r_a$ or $F_a$, see Fig. 1, shows for this female speaker the same overall patterns as for the male speakers in Gobl's study. In Fig. 2 the parameters for the word [behɔ1:a] in stressed position in a sentence are shown for the female and one male speaker. The similarity of the dynamic variation pattern of the source parameters is clear in this picture. The female shows slightly higher values for $r_g$ and higher values for $r_k$ when compared to all three male speakers in Gobl's study. The higher $n_k$ values would indicate a boosting of the lower frequencies of the source spectrum. The $F_a$ value do not differ between these two speakers but, as this female showed the highest $F_a$ values in an earlier study of seven female speakers, Karlsson 1988, we might still expect lower $F_a$ values for female speakers.

Vowels
The vowels are being used to study both intrinsic source differences between vowels of different place of articulation and also hard and loose attack. Very preliminary results indicate that $r_a$, the return time, is relatively longer for a front vowel compared to a back vowel. The isolated vowels, hard attack, start with a considerably lower $r_a$ than the vowels preceded by [h], soft voice onset.

Further results will be given at the conference and will also be illustrated by synthesized examples.

Acknowledgments
This work has been supported in part by grants from the Swedish Board for Technical Development.

References