APPLICATION OF FIR DIGITAL FILTERS TO MODEL HUMAN
AUDITORY PERCEPTION OF SHORT-TIME AMPLITUDE
IN LOW-FREQUENCY RAPID SPECTRUM CHANGE*

Brenda Orser
Department of Linguistics, University of Victoria
P.O.Box 3045, Victoria, B.C. V8W 3P4 Canada

Introduction: FIR (finite impulse response) filters are
applied using CSL,13 in the quantization of the SPL
amplitude component of linguistic resonant pharyngeal
consonants,14 in word-initial CV environments, Fig.1.
Only the onset consonant amplitude responses are modelled,
with FIR transfer functions corresponding to SPL acoustic
responses: glottal excitation (forced response), Fig.2,
separate from the pharyngeal transmission (natural response),
Fig.3. The FIR transfer function (impulse response) system
design has the following properties: pharyngeal transmission
interpreted as odd amplitude/even phase; this is differentiated
from glottal source characteristics: even amplitude/odd
phase. This motivates a DTLTI system.

The objective of the FIR design as applied to speech
transmission is to derive linguistic structure, where
transmission functions correspond to both segmentation in the
auditory system, and active articulatory mechanisms in
speech production. The general issue is the separation of the
system from the source, decoupling vocal-cavity transfer
resonances from vibrations of the vocal cords.1 Basic
design problem is the filter order: specifically, determinants
of the SPL amplitude constant derived from an impulse
which integral to the pharyngeal transfer function will satisfy
a linguistic system’s transmission requirements for the
auditory analysis of pharyngeal resonants. These, as
transients, exhibit an increased and rapid spectrum change of
amplitude in the low-mid frequency region, Fig.3. Also, the
decreasing energy in the glottal excitation contrasts with the
increasing energy exhibited in the transient1 (Figures not
incl.).

Filters: Neither the (temporal) rapidity of the spectral
change, nor the bandwidth are modelled. These instead are
derived as constants, which are linear. An FIR filter
structure is implemented to obtain a linear phase and thus a
constant phase shift, allowing the amplitude response to be
modelled. In an FIR filter the impulse response \( h(n) \) is
limited to a finite number of points.17 The impulse response
is expressed as

\[
(1) \quad h(n) = a_n \quad \text{for } 0 \leq n \leq k, \quad \text{or} \quad (2) \quad h(n) = \sum a_i (n-i)
\]

where \( a_i \) is the number of terms, and function order. The
difference equation for (3) relating the output to the input is

\[
(4) \quad y(n) = \sum_{i=0}^{k} a_i x(n-i) = \sum_{i=0}^{k} h(i) x(n-i)
\]

This describes a nonrecursive realization for the FIR transfer
function.

One significant advantage of the FIR filter function is its
capability in obtaining linear phase (or constant time
delay), derived via expanding the amplitude response in a
sine series. The transfer function also is expanded in sine
terms; the amplitude response can then assume negative
values. This method is possible because at low frequencies
the amplitude response is asymptotic to \( w^k \), with \( k \) odd.19
This is a systemic transmission property.

The FIR filter amplitude response approximating that of
an ideal differentiator is,

\[
(5) \quad A(f) = w \quad (6) \quad B(f) = \frac{1}{1 - 2 \pi f}\]

where \( w = \text{radian frequency (rad/s); } A(f) = \text{amplitude}
response; and } B(f) = \text{phase response, in radians: the phase
associated with the noncausal function combined with the
additional phase due to the added delay.} \]

Segments bearing this type of acoustic feature, and without
phonatory input, are non-contrastive linguistically, and are
attributed to vowel colouring, or, secondary features. The
segments are contrastive only if phonatory features are
added.8-9

Acoustic analysis using cosine terms thus does not
capture the transient’s amplitude change. Also, whereas the
phase shift in cosine terms is ideal, the sine-type expansion
allows for a real linear phase: its phase shift is constant and
at 90°; also as per (6) it is given in radians, etc., and has
all its poles at the origin (and therefore is stable).

The FIR filter order determines the transfer function as a
real function of frequency, thus preferring stability on the
linear phase characteristics. In a DTLTI system the FIR
amplitude response, \( A(f) \), is related as a transfer function
not only to the phase response, \( B(f) \), also to the impulse
response. The latter, approximating a real function of
frequency (i.e., critical band), should satisfy the joint
requirement for transmission: linear phase and constant
time delay. The impulse response is multiplied in a
Hamming window function, with a 12th order bandpass
filter and a low-high cutoff of .04-.13 (Figure not shown).
Thus, both the fundamental and the bandwidth
approximating the second formant are filtered. The
research question then is, how narrow must the main lobe
of the window be so as to satisfy sharp tuning, active
cochlear mechanics, in the auditory system.

Summarily, the filter’s input applied to the waveform
gives a minimalized amplitude response, Fig. 4, D&E
screens (note negative SPL value, in D). The auditory
sound in E is a woody pulse - flatly-tuned and lacking
resonance. This is similar perceptually to the glottal stop,
[?], in which there is minimal laryngeal activity, ideally -
i.e., [-voice], Fig. 2. Alternatively, broadband noise sounds, e.g., [h], can be produced by widening the bandpass cutoff and maintaining even amplitude (filters not shown). The SPL spectrum of either segment precludes mapping to complex auditory responses found in, particularly, the dorsal cochlear nucleus (DCN) responses to linear phase signals in which the amplitude is odd (nonlinear). For example, in damping (Fig. 3, 650–950 Hz); and in sharp tuning resolution found in transient patterns involving significant SPL peaks and troughs. Moreover, the FIR filter’s amplitude response prevents lateral inhibition. Nonetheless, this does not rule out the theory that sharp damping is due to lateral inhibition.\textsuperscript{1} Notably, the missing fundamental alone does not contribute to the lack of amplitude response. The damping function is missing as a factor, prevented by the technical filtering.

**Conclusion:** In applications of the preliminary filter design it is concluded that apparently the auditory system does use SPL amplitude changes of less than 3 Bark, the **critical distance**\textsuperscript{a} across formant bandwidths. But this requires the separation of transmission frequency bands of and for transients, from those of excitation periods. Further, in regard to pharyngeal resonants, in word-initial CV environments, maintaining the amplitude generated by the fundamental neither accounts for nor preserves the apparent transient damping in the second formant region. Whereas FIR filter designs based on source-filter theory require a transfer function that will satisfy formant transitions expanded in terms of bandwidths and the fundamental, the same derivational basis is not present, thus its expansions are not possible, in a DTLTI transmission system specific to pharyngeal resonants. In this, the transfer function is a higher order, a form of transmission, which as a filter can be considered separate from the excitation (source). It should therefore be expanded functionally, as an operation on damping, to thus correspond to segmentation in the auditory system. FIR filtering, with constant phase shifts, allows for this as an operation.

Further research is required to address one of the FIR’s conventional disadvantages: stability of the higher-order filter number and its concomitant increasing time delays in accounting for the observed amplitude response. This could be examined, perhaps by differing filtering and damping in the transmission transfer functions corresponding to delays in the dorsal cochlear nucleus’ complex signal processing.

**References:** *The author wishes to thank D. Wong; Dr.’s. J.H. Esling and B.F. Carlson; and B.C. Dickson. Digital analyses were done in the Linguistics Dept., Phonetics Laboratory, University of Victoria, Canada. Spokane (Interior Salish) tape recordings (1969), provided by B.F. Carlson, were digitized with sampling @ 10k/sec.