

MUSIC AS A METHOD OF IDENTIFYING EMOTIONAL SPEECH

Maartje Schreuder, Laura van Eerten, Dicky Gilbers

HYPOTHESIS

Are differences in emotional speech characterized by different modalities?

Similar to musical modality, we expect to find major third intervals in the intonation contours of cheerful speech and minor thirds in sad speech.

METHOD

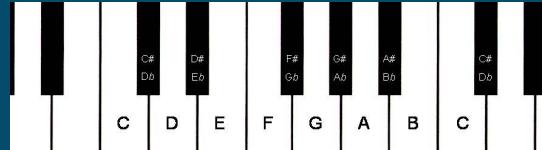
Five professional readers (primary school teachers) each read out four selected stories of the cheerful Tigger and the sad Eeyore (Winnie the Pooh, A.A. Milne) in Dutch.

Relevant passages were extracted and concatenated to twenty files.

The pitch contours (F0) were measured every 10 milliseconds.

We did a cluster analysis of the frequency values which were rounded off to the nearest semitone values, resulting in a semitone power spectrum (cf. Cook 2002, Cook et al. 2004).

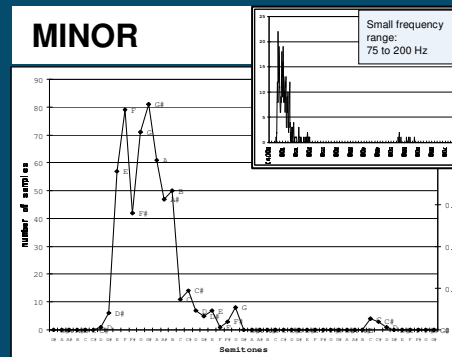
Furthermore, we converted the pitch contours into musical scores.



Minor third: three semitones: C-E b

Major third: four semitones: C-E

MINOR



Semitone power spectrum Eeyore:
interval of three semitones: F-G#

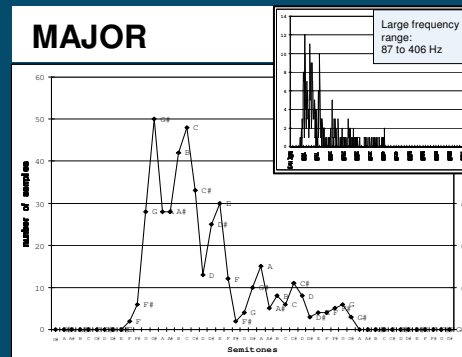
Jejoor voelt zich niet helemaal hoe

Verteld door HJ

Drammatical Ray Jek

Hoe... maak je het? Niet zo eng hoe. Ik voel...
me... al een... he le tijd niet hoe. Wat is er dan mee? Weet je... 't ze ker?
? Wat is... er dan wel? Dat... moet ik zien. Ik ge loof dat... je ge... lijkt hebt.
Dit ver klaart voel. Al... les word me... mi... dui... ch lijk. Deén of
an... chr... moet 'n ge sto... ken heb ben. Zo zijn... ze. Dank je wel, Pooh.
je bent een echte vriend. En dat... kun je niet vanie der een... zeg gen.

MAJOR



Semitone power spectrum Tigger:
intervals of four semitones: G#-C, C-E

Teigetje moet nuzen

Verteld door HJ

Lucky me A. Triad Ajar

Hal lo... ie de reent! Ik heb hem he le maal geen op stop per ver kocht. Nie
jes... ik moesg nie... zen... en toen stond ik net ach ter... joor en toen deed ik,
't was geen op... stop per... ik moest nie zen!

Musical scores of speech fragments incorporate time as a factor in the sequence of tones (ignored in the cluster analyses).

RESULTS

Many Tigger passages have larger intervals than thirds, while many Eeyore files only contain a single peak. In the graphs of these passages the modality cannot be determined.

The major modality is exclusively found in sound files of Tigger in which thirds were observed, whereas the minor modality only appears in sound files of Eeyore.

Although thirds are only found in a minority of our material, there are no counterexamples in the fragments containing thirds.

CONCLUSION

In this pilot study most speech has a neutral modality.

Nevertheless we find a tendency that a sad mood is expressed by means of minor thirds. If thirds are used in cheerful speech, these thirds tend to be major.

We conclude that speech prosody has very similar characteristics as music.

Therefore, music could be a promising method of identifying emotional speech.