PRODUCTION AND PERCEPTION OF PAUSES AND THEIR LINGUISTIC CONTEXT IN READ AND SPONTANEOUS SPEECH IN SWEDISH

Beáta Megyesi & Sofia Gustafson-Capková
Centre for Speech Technology Department of Linguistics
KTH, Stockholm, Sweden Stockholm University, Sweden

Abstract

We investigate the correlation between silent intervals (pauses) and their linguistic context. A considerable attention is given to
a) the syntactic and discourse context in which pauses appear and/or are perceived
b) the characteristics of the discourse structure in terms of produced and perceived pauses.
A case study of three speaking styles with four speakers in each style is carried out.

Method

Investigation of four aspects of communication:
Production of pauses (PA): Automatic detection of pauses in audio data.
Perception of pauses (PP): 18 subjects’ perception of pauses in the same audio data.
Discourse context (TS): 5 subjects annotation of theme shift in the transcriptions of the data.
Syntactic context: The texts were annotated automatically with part-of-speech including morphology, and shallow syntactic analysis (NP, AP, ADVP, PP, NPMAX, etc.).

Speaking styles

• spontaneous speech: elicited dialogues
• professional news reading
• non-professional news reading

Data

Swedish radio news - 920 words each (Sveriges Radio)
* read by 4 professional announcers (2 male, 2 female)
* read by 4 non-professional readers (2 male, 2 female)
Swedish map task dialogue - 920 words (Helgason)
* two dialogues with two participants (1 male, 3 female)

Pause

Silent intervals equal to or longer than 100 ms.; silence incl. breathing and/or swallowing and other natural physical phenomena (but no back channelling).

Theme Shift

No boundary none of the subjects marked a theme shift
Weak boundary 1-2 of the subjects annotated a theme shift
Strong boundary 3-4 of the subjects labeled a theme shift
Extra strong boundary all five subjects agreed on theme shift

RESULTS

PRODUCTION AND PERCEPTION

Professional reading:
The acoustic pauses are perceived (high recall) but a great number of perceived pauses does not have a correlate in acoustic silence (low precision).

Non-professional reading:
Acoustic pauses are perceived (high recall), and perceived pauses have an acoustic silence correlate (high precision).

Dialogue: Few acoustic pauses are perceived (low recall) but many perceived pauses have correlate in silence (high precision).

Recall: the percentage of acoustic pauses that were actually perceived
acoustic & perceived acoustic

Precision: the percentage of perceived pauses that corresponds to silence
acoustic & perceived perceived