

Pause-internal phonetic particles in speech communication

An introduction to the PINTS project

Mikey Elmers, Beeke Muhlack, Raphael Werner, Jürgen Trouvain, Bernd Möbius

Language Science and Technology, Saarland University

www.pauseparticles.org

DFG Deutsche Forschungsgemeinschaft



1 Introduction

Pause-internal phonetic particles (PINTS)

- Breath noises
- Tongue clicks
- Laughing
- Coughing
- Hesitation particles (*uh*, *um*)

Issues

- Incomplete inventory of PINTS
- Lack of unified transcription

Areas of Focus

- Pauses in general
- Pauses in synthetic speech
- Pause modeling and implementation
- Acoustic breath noises with kinematics

2 Goals of this project

- Thoroughly investigate and classify pauses and PINTS
- Improve speech synthesis with human pause patterns
- Evaluate pauses within speech fluency which could influence:
 - non-native speech
 - fluency disorders
 - simultaneous interpretations
 - cognitive diseases (e.g., dementia)
- Determine speaker and language specificity of pausing behaviour and the use of pause-internal particles

3 Research questions & hypotheses

- What are the phonetic characteristics and communicative functions of PINTS and how do they combine in production?
 - vowel quality of hesitation particles (*uh*, *um*) seems to be language-specific [2]
- How do PINTS contribute to the perception of speech?
 - some PINTS shift attention to upcoming material
 - speech fluency is affected by pauses and potentially PINTS
 - breath noises have a strong chance of signaling individuality via idiosyncratic acoustics [3]
- To what degree can synthesized speech benefit from using PINTS?
 - inclusion of breath noises in synthesized speech can enhance processing and perceived naturalness [1]

4 Material

Corpora

- Different languages: German, English, French (and others to be determined)
- Read and spontaneous speech

New Recordings

- Read and spontaneous speech of native German speakers
- Including speech under physical stress (e.g., treadmill running)
- Recordings of audio, Electroglottography (EGG) and Respiratory Inductance Plethysmography (RIP)

Methods

- Production experiments (distribution and acoustic characteristics of pause-internal particles across speakers and languages)
- Perception experiments (acceptance and naturalness of pauses and particles in natural and synthesized speech)

5 Example

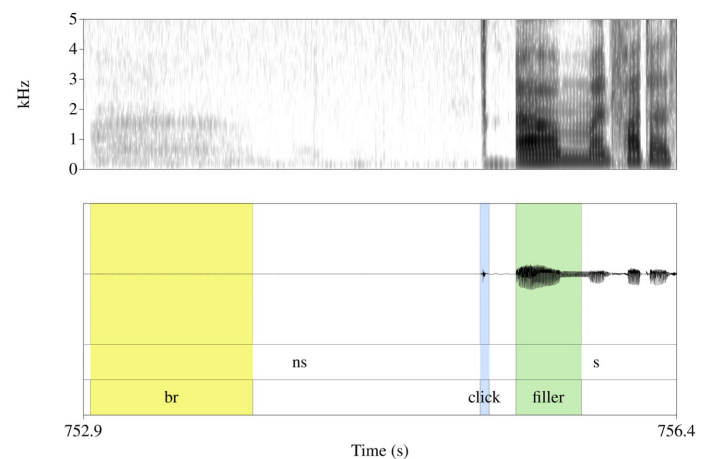


Figure 1: The first tier is subdivided in non-speech (ns) and speech (s), the second tier includes PINTS such as breathing (br), clicks, and fillers (here: *um*).

References

- [1] N. Braunschweiler and L. Chen. "Automatic detection of inhalation breath pauses for improved pause modelling in HMM-TTS". In: *8th ISCA Workshop on Speech Synthesis*. July. 2013, pp. 1–6.
- [2] M. Candea, I. Vasilescu, and M. Adda-Decker. "Inter- and intra-language acoustic analysis of autonomous fillers". In: *DISS 05, Disfluency in Spontaneous Speech Workshop, Sep 2005, Aix-en-Provence, France*. 2008, pp. 47–52.
- [3] J. Trouvain. "Affektäußerungen in Sprachkorpora". In: *Electronic Speech Signal Processing, ESSV 2010, Proceedings of the 21st Conference, Berlin, 8-10 September 2010*. Ed. by H. Mixdorff. Vol. 58. Studentexte zur Sprachkommunikation. TUDpress, Dresden, 2010, pp. 64–70.