

Comparing Annotations of Non-Verbal Vocalisations (NVVs) in Speech Corpora



Jürgen Trouvain & Raphael Werner
 Saarland University, Saarbrücken, Germany
 trouvain|rwerner@lst.uni-saarland.de



1 Introduction

- Non-verbal vocalisations (NVVs) often investigated with experimental data
- Alternative data source: speech corpora
- Unclear: how are NVVs annotated in corpora?
- Aim: to compare annotations of NVVs in various corpora

2 Corpora

- 11 inspected corpora
- German (de) and English (en)
- Spontaneous and scripted speech
- For references see extended abstract

3 Annotation categories

- Only highly frequent types of NVVs
- Laughs and speech-laughs
- Breath noises
- Coughing and throat clearing
- Clicks and lip smacking

4 Overview

y(es) = explicit label found in manual and/or annotation; n(o) = not y

Corpus	Lang.	Mode	Laugh	Speech-laughs	Breath	Cough	Throat clearing	Clicks & lip sm.
ICSI	en	spont.	y	y	y	y	y	n
AMI	en	spont.	y	n	y	y	n	y
Switchboard	en	spont.	y	y	n	n	n	n
Diapix	en	spont.	y	y	y	n	n	y
MapTask	en	spont.	y	n	y	y	n	y
Buckeye	en	spont.	y	y	n	n	n	n
GECO	de	spont.	y	y	n	y	y	n
Lindenstr	de	spont.	y	y	y	y	y	y
KielC.Read	de	scripted	y	y	y	n	y	y
IFCASL	de	scripted	n	n	y	n	n	n
DIRNDL	de	scripted	n	n	n	n	n	n

5 Discussion

- NVVs with a 'Cinderella status'
- Only one corpus showing all NVVs
- Breath noises as most frequent NVV not annotated in all corpora
- Further sub-divisions (e.g. breath and laugh) challenging
- Other studies observe many errors of human annotators (incl. filler particles)
- Inconsistencies probably due to a general lack of awareness of NVVs
- Mid-term goal: more transparency or even a common set of annotation guidelines for NVVs
- Widespread annotation would allow to discover (and ideally understand) NVVs as often subtle phenomena
- Automatic and semi-automatic detection would be very helpful