

## Corpus specifications InterTab

Name of corpus	Tabloid Talk Show Interruption Corpus (subcorpus of the Erlangen Student Corpus Collection)
Abbreviation	InterTab
Brief description of corpus	Transcribed excerpts from tabloid talk shows; focus on conflict talk and interruptions
Name of compiler	Petra Krammer
Type of texts (e.g. genre, language)	Tabloid talk show conversation; American English
Corpus size (in orthographic words)	8765
Number of texts/speakers	20 texts; 3 speakers each (host and two guests; one male, one female)
Length range of texts	Usually two-minute samples
Collection period/date of texts	2014
Release date/version number	11/2019
Source	YouTube clips from <i>The Ricki Lake Show</i> (videos no longer available), <i>The Trisha Goddard Show</i> , <i>The Jeremy Kyle Show</i> (videos no longer available), and <i>The Bill Cunningham Show</i>
Collection method	<ul style="list-style-type: none"> <li>• The corpus consists of twenty excerpts from popular US-American tabloid talk shows and focuses on the interruption behavior of two male and two female talk show hosts.</li> <li>• There are five two-minute excerpts for each talk show host.</li> <li>• Key criteria for the selection of passages: confrontational sequences/ conflict potential.</li> <li>• Communicative situation: one male and one female guest in a conversation with the host. Both guests are expected to contribute to the conversation, but one guest is always mainly addressed by the host and has to defend him-/herself. The samples therefore alternate between male and female principal guests, but not all possibly relevant factors could be considered (e.g. ethnicity).</li> <li>• Shows: <i>The Trisha Goddard Show US</i>, <i>The Ricki Lake Show</i>, <i>The Jeremy Kyle Show US</i> and <i>The Bill Cunningham Show</i>.</li> </ul>
Compilation and annotation conventions (e.g. tokenisation, lemmatisation, tagging, syntactic parsing)	<p>Based on Hutchby, I. &amp; R. Wooffitt (2008). <i>Conversation Analysis</i>. 2nd edn. Cambridge: Polity Press. (pp. 73-87.)</p> <p>(0.5) the number in brackets indicates a time gap in tenths of a second          (.) very short gap          = ‘latching’ between utterances or continuation of a turn across intervening lines of transcript          [ ] square brackets between adjacent lines of concurrent speech indicate the onset and end of overlapping talk          [[ ]] simultaneous start of speakers’ turns          hh h indicates out-breath; the more h’s, the longer the breath          (( )) a description enclosed in a double bracket indicates an activity, e.g. ((laughing))          soun- a dash indicates the sharp cut-off of the prior word or sound/ stammering quality          sou:::nd colons indicate lengthening of the preceding sound; the more colons, the greater the extent          ! animated or emphatic tone          ( ) empty parentheses indicate the presence of an unclear fragment on the tape          XXXXX unintelligible word/phrase/sentence          AAAAA irrelevant ongoing talk          CAPITALS capitalisation marks speech as noticeably louder than surrounding speech</p>

Sociolinguistic description of speakers (age, sex, native language, education etc.)	Various
Research project for which corpus was compiled	<i>Gender as a determinant of the interrupting behavior of male and female hosts of American tabloid talk shows?</i> Coursework in the Hauptseminar <i>Linguistic variation and change in present-day English</i> , Summer semester 2014, FAU (Friedrich-Alexander-Universität Erlangen-Nürnberg).
Name of supervisor	Christina Sanchez-Stockhammer
Contact person/contact details	Christina Sanchez-Stockhammer christina.sanchez@lmu.de
Project homepage	<a href="http://www.erlanger-linguistik-online.uni-erlangen.de/projekte/escs.shtml">http://www.erlanger-linguistik-online.uni-erlangen.de/projekte/escs.shtml</a>
Disclaimer	You are free to use the corpora provided here for academic research. Please inform the authors of the corpora about any research projects using their material.
Citation	Krammer, Petra. 2014. Tabloid Talk Show Interruption Corpus (InterTab). <i>Erlangen Student Corpus Collection (ESCC)</i> . <a href="http://www.erlanger-linguistik-online.uni-erlangen.de/projekte/escs.shtml">http://www.erlanger-linguistik-online.uni-erlangen.de/projekte/escs.shtml</a>