

Corpus specifications FCC

Name of corpus	Football Commentary Corpus (subcorpus of the Erlangen Student Corpus Collection)
Abbreviation	FCC
Brief description of corpus	live soccer game commentary
Name of compiler	Daniel Meinl
Type of texts (e.g. genre, language)	live sports game commentary, British English
Corpus size (in orthographic words)	7188
Number of texts/speakers	6 text, 2 speakers per text, 4 in total (sports commentators)
Length range of texts	10-minute samples
Collection period/date of texts	broadcast November 2010 – April 2014, collected September 2014
Release date/version number	11/2019
Source	YouTube clips, SkySports
Compilation and annotation conventions (e.g. tokenisation, lemmatisation, tagging, syntactic parsing)	<p>---- = unintelligible/omission pauses: (.) 1-2 seconds (-) about 5 seconds () about 10 seconds</p> <p><ã> in <i>Coentrão</i> was replaced with <a> <ë> in <i>Raphaël</i> was replaced with <e> <Ö> in <i>Özil</i> was replaced with <O> <i> in <i>Higuaín</i> was replaced with <i> <é> in <i>Olés</i> was replaced with <e></p> <p>HTML symbol masking: < = &lt > = &gt</p>
Sociolinguistic description of speakers (age, sex, native language, education etc.)	Kevin Keatings, age unknown, male, British Terry Gibson, 51, male, British Rob Palmer, age unknown, male, British Gerry Armstrong, 59, male, British (Northern Ireland)
Research project for which corpus was compiled	<i>How conversational is the turn-taking in football commentary with two speakers?</i> Coursework in the Proseminar <i>Meaning and communication</i> . Summer semester 2014. Friedrich-Alexander-Universität Erlangen-Nürnberg.
Name of supervisor	Christina Sanchez-Stockhammer
Additional information	-
Contact person/contact details	Christina Sanchez-Stockhammer christina.sanchez@lmu.de
Project homepage	http://www.erlanger-linguistik-online.uni-erlangen.de/projekte/escc.shtml
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	Meinl, Daniel. 2014. Football Commentary Corpus (FCC). <i>Erlangen Student Corpus Collection (ESCC)</i> . http://www.erlanger-linguistik-online.uni-erlangen.de/projekte/escc.shtml