

Corpus specifications BBSCC

Name of corpus	Blood Bowl Stream Commentary Corpus (subcorpus of the Erlangen Student Corpus Collection)
Abbreviation	BBSCC
Brief description of corpus	live video game commentary streamed on Twitch.tv
Name of compiler	Heiko Ermer
Type of texts (e.g. genre, language)	live video game commentary, American and British English
Corpus size (in orthographic words)	6079
Number of texts/speakers	7 texts / 1 or 2 speakers (broadcasters) + users from chat
Length range of texts	1-8 minute samples
Collection period/date of texts	January to February 2017
Release date	11/2019
Source	live stream, www.twitch.tv (no longer available)
Collection method	<ul style="list-style-type: none"> - the corpus consists of the farewell section as well as one randomly selected passage from the broadcast - the farewell sections are about 1 minute long, the passages from the broadcast approximately 8 minutes - communicative situation: one or two broadcasters interact with each other and their viewers through spoken language. Viewers address each other and the broadcasters via chat-room messages - games played: Blood Bowl 2
Compilation and annotation conventions (e.g. tokenisation, lemmatisation, tagging, syntactic parsing)	HTML symbol masking: - < = < - > = > - & = &
Sociolinguistic description of speakers (age, sex, native language, education etc.)	Crendor: 26 years, male, native English speaker (American) Strippin: 28 years, male, native English speaker (British)
Research project for which corpus was compiled	<i>A Linguistic Analysis of Video Game Commentary</i> . BA thesis. 2017. FAU (Friedrich-Alexander-Universität Erlangen-Nürnberg)
Name of supervisor	Christina Sanchez-Stockhammer
Acknowledgements	The transcripts were published with kind permission from Crendor.
Additional information/Comments	annotation of sample linguistic features in the BA thesis - <i>blodge</i> : a unit having both the Block and Dodge trait - <i>tv</i> : an abbreviation of team value - <i>ty</i> : an abbreviation of “Thank you” - <i>GG</i> : an abbreviation of “good game” - <i>inb4</i> : an abbreviation of “in before” - <i>U+1F44C</i> : unicode “ok hand sign” - <i>Surf's up</i> : shortening of: the surf is up - <i>Nuffle</i> : a fictional deity in the Blood Bowl universe
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	Ermer, Heiko. 2017. Blood Bowl Stream Commentary Corpus (BBSCC). <i>Erlangen Student Corpus Collection (ESCC)</i> . http://www.erlanger-linguistik-online.uni-erlangen.de/projekte/escc.shtml >