

Swiss SMS Corpus

<https://sms.linguistik.uzh.ch/>

Sammlung von 25.947 SMS, die in dem Zeitraum 2009-2011 von 2.784 Teilnehmern gesammelt wurden.

Sprache	Schweizerdeutsch, (Standard)-Deutsch, Französisch, Italienisch, Rätoromanisch
Sprachliche Realisierung	schriftlich
Umfang	~650'000 tokens
Medium	SMS Nachrichten
Geographischer Ursprung	Schweiz
Zeitliche Einordnung	2009-2011
Form der Daten	Daten wurden für zwei Korpora unterschiedlich aufbereitet: 'ANNIS' unterteilt die SMS nach Sprache und hat eine ausführliche Annotation (PoS-Tags), 'SMS Navigator' enthält Metadaten zu zwei Drittel der Daten und hat eine benutzerfreundliche Suchfunktion
Format	XML
Annotation	Part of Speech Tags (ANNIS), soziodemographische Metadaten (SMS Navigator)
Quelle /Herausgeber	Elisabeth Stark, Simone Ueberwasser, Beni Ruef; Swiss National Science Foundation
Nutzungsvoraussetzungen	Zugang frei, Anmeldung erforderlich
Link	https://sms.linguistik.uzh.ch/Main/WebHome
Literatur	Stark, Elisabeth; Ueberwasser, Simone; Ruef, Beni (2009-2015). Swiss SMS Corpus. University of Zurich.