Hans-Jörg Schmid, Quirin Würschinger, Melanie Keller
and Ursula Lenker

Battling for semantic territory across social networks. The case of *Anglo-Saxon* on Twitter

Abstract: In Present-Day English, the term *Anglo-Saxon* is used with three dominant meanings, which have been labeled "historical/pre-Conquest", "ethno-racial" and "politico-cultural" uses (cf. Wilton 2019). From at least the middle of the 19th century, the second sense has been politically appropriated to convey the racial notion of white supremacy. Recently, a fierce conceptual and socio-political controversy over the meaning and implications of the term *Anglo-Saxon* has spilled over into academia, ultimately causing the vote of the members of the *International Society of Anglo-Saxonists* (*ISAS*) to change its name to *International Society for the Study of Early Medieval England* (*ISSEME*).

We discuss this development as a paradigm case of controversies over the conceptual territory associated with a contested term. Using the Entrenchment-and-Conventionalization Model (Schmid 2020) as a theoretical framework, we analyze a large collection of Twitter posts (n ~ 510,000) with a view on the conventionalization and entrenchment processes involved in the conceptual and political controversy.

Keywords: semantic contestation, semantic variation, semantic change, social variation, Twitter, corpus linguistics, collocation analysis, Social Network Analysis (SNA)

1 Introduction: The dominant meanings of *Anglo-Saxon* in Present-Day Englishes

The term *Anglo-Saxon* (n. and adj.) is used with three dominant meanings in Present-Day English: It can refer to (1) England and its English-speaking inhabitants

Hans-Jörg Schmid, Ludwig-Maximilians-Universität München, hans-joerg.schmid@lmu.de
Quirin Würschinger, Ludwig-Maximilians-Universität München, q.wuerschinger@lmu.de
Melanie Keller, Ludwig-Maximilians-Universität München, Melanie.Keller@anglistik.uni-muenchen.de
Ursula Lenker, Ludwig-Maximilians-Universität München, Ursula.Lenker@lmu.de

before the Norman Conquest (*OED* s.v. n. and adj. 1), (2) people of English (or British) heritage or descent, or, more generally, of Germanic origin, and hence: white and English-speaking (*OED* s.v. n. and adj. 2), and (3) the predominantly English-speaking countries and their cultures, politics and economics. These three broad meanings have been labeled (1) "pre-Conquest" or "historical uses", (2) "ethno-racial uses" and (3) "politico-cultural uses" (following Wilton 2019). For examples from our Twitter data, see examples (1)-(3):

(1)   Two Anglo-Saxon shoulder clasps both dating to around the early 7th century which put on full display one field the Anglo-Saxons as a whole excelled in, metalworking.

(2)   (a) I'm peeling like a molting iguana. Thank you, Anglo-Saxon heritage!"
       (b) ya'll so white and anglo saxon and shit

(3)   professor paolo paolini points out web 2.0 is a largely monocultural (anglo-saxon), monolingual (english) phenomenon. why is that exactly?"

The 'historical/pre-Conquest' sense of *Anglo-Saxon* has a long academic tradition in historical studies, and consequently in the British educational system, in library cataloguing, etc. Less prominent in the minds of most medievalists and linguists, at least until recently, the term has a parallel history, about 150 years long, of being used to denote people of English (or British) descent and their culture(s) and particularly to white English-speaking people of English heritage (i.e. the "ethno-racial uses"). This last sense has lately given rise to a fierce academic and political dispute.

The debate revolves around the 'ethno-racial' sense of the term *Anglo-Saxon*, which may seem unrelated to its 'historical/pre-Conquest' sense. A closer look at its history, however, reveals nationalist or racist mentalities in the conscious political appropriation of the term from the early nineteenth century onwards, when it was used to endorse white supremacy in both Europe and the United States (see Utz 2019 and, for the US e.g., Horsman 1985, Dockray-Miller 2017). Some recent in- and out-group critics of the field of 'Anglo-Saxon studies' even feel that this interdependence of nationalism and medieval studies deems the field 'by etymology and origin' racist or, at least, that it causes misconceptions of the field of Anglo-Saxon studies as being inherently (i.e. by virtue of the term *Anglo-Saxon*) nationalist or racist. US language users in particular seem to connect the term instantaneously to what could "be summed up in the acronym WASP – White, Anglo-Saxon and Protestant – a code for racial purity that white supremacists and neo-Nazis have embraced" (Wood 2019). Others, by contrast, stress the utility of the term *Anglo-Saxon* for precisely referring to a particular period in the history of England and its people(s) and culture(s). This encompasses much more than an 'Anglo-Saxon ethnicity' (i.e. senses (2) and (3)), since in contemporary historical studies, the meaning of *Anglo-Saxon* is not limited to

the study of ethnic Angles and Saxons, but to the specific character of the multicultural and multilingual period before the Normal Conquest in a distinct mix of 'English' (Angles, Saxons, Jutes, Frisians, if we want to stick to these established terms), Celtic, Romano-Celtic, Latin and Scandinavian, etc. cultures.

At some points in the recent academic and non-academic discussions, these two positions were so fiercely defended that we saw a battle of arguments over the semantics of *Anglo-Saxon*, in which misunderstandings and accusations often seemed to gain dominance.[1] One reason for the fierceness of the debate may be grounded in the major differences of usage between the contemporary varieties of English: Based on a comprehensive corpus analysis covering the period from Old English to 2017,[2] Wilton (2019) finds that the 'racial' sense has become the predominant one in contemporary North American speech (US and Canada), with some 66 per cent of uses (and only 22 per cent for the 'historical use'). In contemporary British English, the numbers are reversed, with about 75 per cent for the 'historical use' (and even 87 per cent in academic registers; see Wilton 2019: 26, 30).[3]

Using the controversy over *Anglo-Saxon* as a case study, we hope to contribute to a better understanding of the dynamics of battles over territory in conceptual and semantic space taking place on social media. We add to Wilton's account in two ways. We analyze a large dataset from Twitter reaching up to May 2020, because the social media platform plays an increasingly important role as an arena for political controversies, including the one examined here. In contrast to other types of data, the material from Twitter allows us to analyze the social dynamics of the process. This is an important aspect in our general endeavor to unveil the linguistic, social and cognitive processes involved in the conceptual and political controversy.

2 Material and theoretical background: Twitter data and the EC-Model

We base our investigation on a comprehensive collection of tweets on the query term *Anglo-Saxon* since the start of Twitter in December 2006. We have created

---

1 For different perspectives on the debate in what was the International Society of Anglo-Saxonists see, e.g., Utz 2019, Wilton 2019, Wood 2019 and www.fmass.eu/uploads/pdf/responsible_use_of%20the%20term%20_Anglo-Saxon.pdf.
2 Among them COHA, the Hansard Corpus, COCA, the Strathy Corpus of Canadian English, NOW and the BNC.
3 In British English, the 25 per cent of examples for sense (2) are (also) found in the context of an ethnic identity label distinguishing the English from the Welsh, Irish or Scottish.

this dataset using the Python library *twint*,[4] which relies on Twitter's Advanced Search feature. We filter the collected tweets using several post-processing steps in *R*. We limit the corpus to tweets which contain the term *Anglo-Saxon* or a range of orthographic variants[5], and we discard hashtag uses and duplicate tweets. The resulting dataset contains 506,576 tweets, with the first tweet dating from 14 December 2006 and the last tweet from 8 May 2020.

We analyze the data and interpret our findings from the perspective of the Entrenchment-and-Conventionalization Model (EC-Model, Schmid 2020). We consider this dynamic usage-based socio-cognitive model particularly suited for endeavors such as the present one because it offers a systematic account of the usage-driven interplay between the conventionalization of meanings, the social dynamics of groups and the cognitive representations in the minds of individual speakers. Essentially, the model claims that language is a two-fold feedback system in which usage (represented in the center of Figure 1) both drives and is driven by social processes in a community (conventionalization feedback cycle) and by cognitive processes in the minds of individual speakers (entrenchment feedback cycle). The systematic separation of usage, communal conventionalization and individual entrenchment will be helpful for understanding the dynamics of the conflict investigated here.
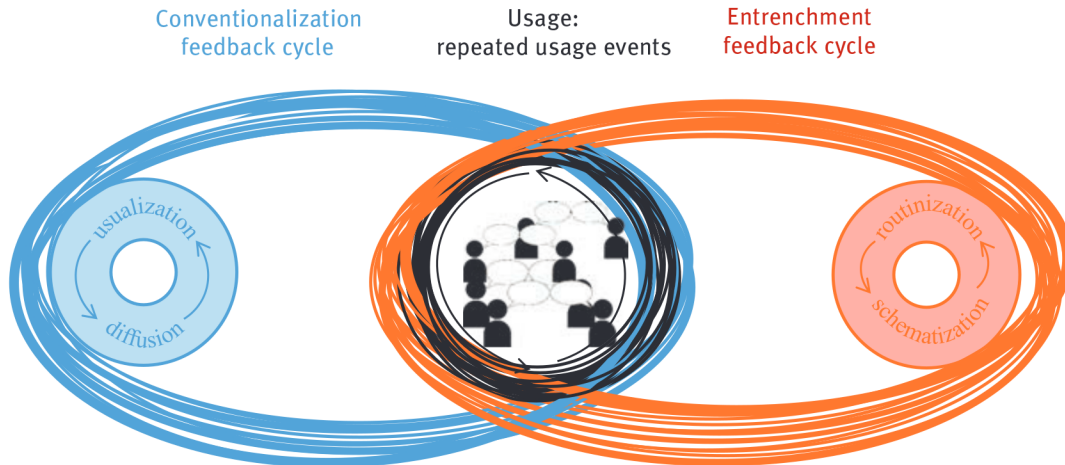

Figure 1: A simplified version of the Entrenchment-and-Conventionalization Model (adapted from Schmid 2020: 4)

---

4 https://pypi.org/project/twint/
5 We use the R package stringr to perform regular expression matching using the pattern ^[^#]*(A|a)nglo(-|\\w|\\b|\\s)?(S|s)axon

3 Research questions and methods

Picking up the colors from Figure 1, Table 1 lists the core components of this model and the relevant questions we aim to answer, as well as the corresponding foci of analysis and the methods used.

Table 1: Design of the study based on the EC-Model (Schmid 2020)

		Research question		Focus of analysis	Method
USAGE		How and with what motivation is *Anglo-Saxon* used in the tweets?		Speakers' communicative goals and social forces affecting their linguistic choices	Induction of communicative goals and forces from manual inspection
CONVENTIONALIZATION		To what extent are the different meanings of *Anglo-Saxon* usualized, how are the meanings related to social parameters of users and social networks, and how has this changed over time?	USUALIZATION	Usualization of form-meaning regularities, syntagmatic regularities, contextual and social regularities	Induction of denotative, connotative and social meanings from manual inspection and collocation analysis; analysis of change in collocations over time
			DIFFUSION	Geographical and social diffusion of regularities	Induction of individuals and social groups involved in usage by means of Social Network Analysis
ENTRENCHMENT		To what extent have individual speakers formed habits with regard to their use of *Anglo-Saxon*?	ROUTINIZATION	Routinization of symbolic, syntagmatic, paradigmatic, and pragmatic associations	Induction from manual inspection and pattern analysis of data of selected influential agents

4 Usage: speakers' communicative goals and social motivations

Our detailed manual investigation of the Twitter data largely confirms many of the results of Wilton (2019). Following Schmid (2020), we describe the four dominant utterance types by specifying speakers' communicative goals and social motivations:
1.   'Historical/pre-Conquest' use: referring to the period, people(s) and culture(s) of pre-Conquest England without any noticeable marking of stance.

2. 'Ethno-racial' use: referring, usually with either a positive or negative stance, to people of white, English-speaking heritage, motivated by the wish to mark either in-group or out-group identity, to disparage others, claim power and question or reinforce social order.
3. 'Politico-cultural' use: referring to commonalities attributed to the US and the UK and possibly other English-speaking countries, often from an external perspective.
4. 'Metalinguistic' use: discussing either the term *Anglo-Saxon* itself or related words and expressions from a metalinguistic perspective with regard to their meanings, appropriateness, etymology, etc.

The four utterance types are illustrated by examples (4) to (7) respectively:

(4) On another note. I have recently been sent a very interesting paper on Anglo Saxons coins. It's going to be up on... [URL]
(5) What the hell is that Anglo Saxon bitch, Yvette Cooper up to? [URL]
(6) [@USER] it's clear you're trapped in an Anglo-saxon system. Unaware of education in continental Europe. Please read up
(7) Frodo means wise in Norse/Anglo Saxon #fyi.

5 Conventionalization: how conventions are sustained and renewed

In the EC-Model, conventionalization is conceived of as an ongoing dynamic process. It is defined as the continuous sustaining and renewing of conventions, understood as regularities of behavior. These regularities relate to six different dimensions of conformity:
– onomasiological conformity of meaning/intention-form correspondences, e.g. referring to a person as *Anglo-Saxon* or *WASP* or *white* etc.;
– semasiological conformity of form-meaning correspondences, e.g. associating the form *Anglo-Saxon* with the meaning 'pre-Conquest' or 'ethno-racial' or 'politico-cultural';
– syntagmatic conformity of sequential patterns, e.g. associating *Anglo-Saxon* with typical collocates such *England* or *protestant* or *culture* etc.;
– cotextual conformity of correspondence of linguistic choices to genres and text-types, e.g. using the term *Anglo-Saxon* in an academic paper, a tweet, an opinion piece in a newspaper, etc.;
– contextual conformity of correspondences of linguistic choices to situations and goals, e.g. using the term in an academic meeting, when retweeting, or when discussing politics in a bar;

– social conformity of correspondences of linguistic choices to social aspects of users, e.g. uses of the term *Anglo-Saxon* by members of different social groups or communities of practice.

Although all six dimensions of conformity constantly interact, the first three are primarily established and sustained by the process of usualization, and the second three dimensions by the process of diffusion. The following discussion is divided into two parts focusing on usualization and diffusion, both of which include a diachronic perspective to integrate the dynamic nature of convention-alization.

5.1 Usualization (and its relation to diffusion): how meanings are negotiated

Our main source of evidence for the observation of usualization comes from the analysis of collocations. Based on Firth's (1957: 11) dictum that "you shall know a word by the company it keeps", we interpret the findings of the collocation analysis not only in terms of syntagmatic, but also onomasiological and semasiological conformity, and further include our assessment of relations to situational and social conformity. All interpretations of collocational evidence are supported by the close manual inspection of a very large number of concordance lines.

In order to produce a longitudinal diachronic study of collocations, we divided the material into batches of 20,000 tweets and used Antconc (Anthony 2014) to identify significant collocates for each period.[6] For maximum transparency of results, we opted for a descriptive report of consistent, transient and initiating collocates (see McEnery, Brezina and Baker 2019), i.e. collocates that are

– consistent, i.e. associated with *Anglo-Saxon* over the whole 14-year period from 2006 to 2020,
– transient, i.e. appear for some time and then disappear, or
– initiating, i.e. begin to be associated and stay in place until the end of the period under investigation.

---

6 The following settings were used for the collocation analysis in Antconc: The search term for concordancing and collocation analysis was "[^\#][A,a]nglo[\-,\w,\b,\s]?[S,s]axon" (regular expression), retrieving variants such as Anglo-Saxon, Anglo-saxon, anglo-saxon, Anglosaxon but excluding hashtag uses, i.e. #anglosaxon. The settings for collocation analysis were: measure of association: t-score; span: 4L to 4R; minimum collocate frequency 100; no stop list was used, function words and other irrelevant collocates, e.g. http, com, url, were deleted manually.