



# Almans and Kanaks: Inter-cultural Stereotypes in a German Twitter Corpus

ECKHARD BICK

Department of Language and Communication, University of Southern Denmark

## ABSTRACT

This paper examines inter-cultural stereotypes in a large German Twitter Corpus (2 billion words). The main focus is on hetero- and auto-stereotypes found within the Turkish and Arab immigrant community in Germany, and since much of the relevant German-language data is influenced by the Kietzdeutsch sociolect, the term *Alman* (Turkish for 'German') is used as a window into the corpus. The term is contrasted with the German slur *Kana(c)ke* ('foreigner') and both investigated for associated stereotypes in both qualitative and quantitative ways, using concordances and relevant linguistic patterns (e.g. generalization and othering), as well as relative frequencies, sentiment analysis and word vector distances based on emoji classification. In connection with negative stereotyping, we discuss examples of "counter hate speech" directed by (part of) the immigrant minority at the German majority. Finally, we investigate the role of humour and jokes as vehicles of stereotypes.

## KEYWORDS

Social Media, Hate Speech, Corpus Linguistics, Immigrant Auto- and Heterostereotypes, Kiezdeutsch, Slur Compounding, Othering, Alman Jokes

## 1. Introduction

In the wake of accelerated demographic change and multiculturalization of Western societies, Hate Speech (HS) against immigrant and refugee minorities (e.g. Foxman & Wolf 2013) has become a growing concern, not least in social media, where communication is often less premeditated, more personal and certainly less edited than in mainstream media. Besides abusive and aggressive speech and demeaning metaphors, ethnic and religious stereotypes are an important vehicle for the generalization and othering processes involved in the expression of HS. However, most research in the area has focused on stereotypes and HS manifesting in a host society against immigrant minorities rather than the other way around (e.g. Jaki & De Smedt 2019), sometimes making the minority target a condition sine qua non. Thus, Human Rights Watch defines HS as "any form of expression regarded as offensive to racial, ethnic and religious groups and other discrete minorities, and to women" (Bell 2009). For a corpus-linguistic approach, the problem is compounded by the lack of data, since ethnic minorities are less likely to participate in the online discourse due to barriers such as language, access, or education. In

this paper, we try to address this sparse data problem by using the Turkish slur demonym *Alman* ('German'), in a German Twitter corpus, as a corpus-linguistic window into stereotyping processes and HS directed from the immigrant minority at the (German) majority. The 2 billion word corpus<sup>1</sup> was compiled as a monitor corpus in the context of a 3-year HS project (XPEROHS, Baumgarten et al. 2019) and annotated with morphosyntactic and semantic-class tags, dependency structure and verb frames (Bick 2020).

## 2. The term *Alman*

### 2.1. Meaning

The word *Alman* was introduced into German by the Turkish minority in Germany, consisting of descendants of the *Gastarbeiter* invited into the country to fill labor vacancies. The term is often used in a derogatory context but is not necessarily a slur on its own. Nor is it always used for ethnic Germans (1), and can also mean a *Deutschländer*, a person of Turkish descent not any more perceived as properly Turkish. In the latter sense, *Alman* is always negative (2).

- (1) *Almans sind ja nicht unbedingt Deutsche, sondern halt so richtige almans.*  
(‘Of course, Almans are not all Germans, but precisely typical Almans’)
- (2) *Ich kann nichts dafür, dass meine Großeltern 50 Jahre früher als Flüchtlinge gekommen sind. Deswegen muss ich mich nicht als Alman beschimpfen lassen.*  
(‘It isn’t my fault that my grandparents came 50 years earlier than the refugees. That’s no excuse for insulting me as an Alman’)

The word has become part of the vocabulary of *Kiezdeutsch*<sup>2</sup> (Freywald et al. 2011), a youth sociolect spoken in immigrant neighborhoods by Turkish, Arab and other ethnicities as well as young Germans identifying with the same community. Over the last decade or two, *Alman* has also become an important vehicle of irony, including self-targeting irony, when venting stereotypes about Germans (*Almans*) and immigrants (*Kanaks*), giving rise to a veritable genre of Alman jokes with a fixed narrative pattern. In this setting, and as an auto-stereotype, even the word *Kanacke*<sup>3</sup>, otherwise a strong slur and quite derogatory, can morph into ironic or neutral territory<sup>4</sup>. Though the embedding sociolect is linked to a negative out-group picture (Wiese 2011), and activation of associated negative stereotypes can compromise school performance of in-group members (Martiny et al. 2014), the use of *Alman* and *Kanak* in our corpus, as well as the stereotypes linked to them, suggest at least a counter current, with increased self-assertion and Germans (*Almans*) in the role of outsider, suggesting an inverted, and sometimes aggressive, ingroup-outgroup discourse.

1 The corpus is a monitor corpus. 2 billion was the size at the time of data extraction - the current size is about 2.7 billion.

2 Also called *Ghettoslang*, *Kanak Sprach*, *Kanakisch*, *Türkendeutsch* or *Ausländerdeutsch*

3 The usage difference is often visible at the spelling level, with *Kanacke* becoming *Kanak* in *Kiezdeutsch*.

4 The ironic potential of the term is stressed by Mandel (2008, p. 2), who sees the term as having been “transformed into an indirect moniker of subaltern power”, and attributes the shift to the work of writer-rapper Feridun Zaimoğlu.

## 2.2. Morphology

There are a number of morphological variants of *Alman* (which is relevant for corpus searches), some more "Turkish" (*almani*, *amanci*), some more "German" (*Almanin* ['she-Alman'], *Almann* (German 'nn'), *Almanen* (German plural), *Almaner/in*, *Almanese* or even the iconically German<sup>5</sup> *Almaner\*innen*) or even "Romance" (*almano*, *almana*). In addition, the term supports derivation, affixing and compounding. Thus, the Turkish toponym *Almanya* ['Germany'] gets the synonyms *Almanland*, *Almanien*, *Almansland*, *Almanesien* and the islamization word play *Almanistan*. Similarly, there are adjectives for prototypical Alman behavior (*almanhaft*, *almanesisch*, *almanisch*, *almanisch*, *almanisch*, *almanig*, *almanisiert*, *almanesk*, *almanmässig*, *almanisch*, *almanesque*), nouns denoting "almanhood" (*Almanismus*, *Almantum*, *Almanistik*, *Almankeit*, *Almanität*, *Almanness*) and even verbs (*almanisieren*). At least at the morphological level, German grammar does apply, as can be seen from verb forms such as *angealman* ('having been insulted in an Alman way') or *rumzualmanen* ('doing or saying meaningless things in an Alman way').

## 2.3. Compounding

The corpus also exhibits a great deal of ordinary German productive compounding, e.g. the neutral *Alman-Mädchen/Kinder/Familie/Oma/Opi* (*Alman*+family term), the less neutral *Alman-Weib* ('Alman [vulgar] woman') and the derogative *Almanscheisse/-kacke/-shit* or *Scheiß-/Kack-/Drecksalman* ('dirt[ly]'/ 'shit[ty]' Alman), as well as the superlatives *Über-/Uberalman*, *Superalman*, *Ultraalman*, *Premiumalman*, *Oberalman* ('Alman-in-chief'), *Alphaalman*, *Vollblutalman* ('thoroughbred Alman') and *Protoalman*.

The range of *Alman*-compounds found in the corpus also illustrates that demeaning and hateful speech is by no means a one-way phenomenon directed at immigrants by Germans, but also works nicely the other way. The most common is *Hurensohn* ('son-of-prostitute'), which is used about Germans both on its own, and as first and second elements (*Almanhurensohn*, *Hurensohnalman*), and the related *Bastard* (*Almanbastard*, *Almanpicos* [*piç* - Turkish: 'bastard']). Other terms have a sexual connotation (*Almanfotze* ['cunt'], *Almanschwuchtel* ['she-gay']) or are based on animals perceived as dirty (*Almanköter/-hund* ['dog'], *Almansau* ['pig'], *Almanratte* ['rat']). Specific stereotypes about Germans enter into *Almannazi/Nazialman* and *Faschoalman* (history: 'nazis'), *Kartoffelalman/Almankartoffeln* (food: 'potatoes'), *Schweinefresser-Alman* (forbidden food: 'pig eater') and *Lauchalman/Almanlauch* (body: 'leek', i.e. long & thin). More generically, compounds about diminished manhood (*Almancuck* - Turkish: 'sissy') or intellect (*Almandepp* ['idiot'], *Vollalman* ['complete alman'], *Alman-Otto*).

More neutrally, compounds can denote things typically Alman, e.g. *Almanstäbchen* ('cash desk divider'), *Almanwurst* ('sausage with pig meat'), *Almandudler* (word play about a drink), and societal conditions, e.g. *Quotenalman* (minimum ethnic representation in schools). Last not least the term is used in humoristic word creation, e.g. *Bioalman* (all Germans are eco-freaks<sup>6</sup>), *Alpenalman/Bergalman* ('Austrian') and *Alman-o-meter* or word plays like *Almankalypse* now. At

5 Various orthographic strategies of gender-neutralisation have gained immense weight in official German writing, among them the famous *Gendersternchen* ('gender asterisk') introducing an added female ending.

6 The prefix *Bio-* means 'Eco-' and is frequently used in connection with, e.g., consumer products.

a reflective meta-level, we find words like *Almanbashing/-shaming*, *Schuld-kult-Alman* (German historical bad conscience) and (3) *Almansplaining*:

- (3) *find wichtig, dass jetzt almans Menschen mit migrationshintergrund erklären, was rassismus ist*  
(‘I think it important that Almans now want to explain racism to immigrants’)

### 3. Stereotypes

#### 3.1. Corpus evidence

Even though *Alman* is a more neutral term than a classical slur like *nigger*, the term is laden with stereotypes. In fact, in our Twitter data, the majority of occurrences can be linked to stereotyping, and in many cases stereotyping is quite explicit, as with the use of *Almanding* (‘Alman thing’) in examples (4) (Alman are hobby politicians) and (5) (Almans are car freaks)

- (4) *Sich über irgend ne Mülle beim Frühstückfernsehen aufzuregen ist auch so ein Almanding*  
(‘Getting upset about some woman\_slang in breakfast tv is also one of these Alman things’)
- (5) *Ist wahrscheinlich krankes almanding sowas zu sagen aber ich bin begeistert von südafrikas fahrbahnmarkierungen (sic!)*  
(‘It’s probably a sick Alman thing to say this, but I’m so impressed by South Africa’s motorway lane markings’)

Since stereotypes are by definition simplistic, and often used for othering or delimiting group identities, they thrive in a contrastive setting. This can be implicit like in the ironical (6), where the stereotype is that Almans don’t know how to party, or explicit as in (7), where Almans are stereotyped as non-macho eco-Germans.

- (6) *willst du hochzeit oder almanmässig mit 20 personen*  
(‘Do you want a wedding or Alman-style with 20 people’)
- (7) *Nigas und kanacken geben sich anabolen und almans kaufen sich haferflocken*  
(‘Nigas and kanacks take anabolic steroids and almans buy oat cereals’)

Stereotyping is often linked to certain linguistic constructions. For non-contrastive stereotyping this can be generalization markers (plural/*these/all*) that will be further explained in chapter 4, or *wenn*-constructions, e.g. (8) where Almans are depicted as culturally insensitive. For contrastive stereotyping, joke formulas are common, a topic that will be treated in more detail in chapter 5. Example (9), for instance, humoristically contrasts Alman hyper-rationality with Kanak gut feeling.

- (8) *wenn Almans Loriot Witze rezitieren und davon ausgehen man kennt die auch*  
(‘when Almans recite Loriot jokes and think you know them, too’)
- (9) *Sportwetten - Almans: betreiben wochenlange Analyse, recherchieren auf jeder Seite, gucken wann die Spieler zuletzt Geschlechtsverkehr hatten | Kanaken: Bruder vallah fener die können das Ding machen*  
(‘Sports betting - Almans: week-long analysis, research all aspects, when did the players last have intercourse | Kanaks: Bro vallah fener (FC<sup>7</sup>) could swing it’)

It is interesting to note that when used by immigrants for positive or ironic auto-stereotyping, the terms *Kana(c)k(e)* and *Niga* are more neutral, and not the slurs they would have been, had a German used them.

In their study of immigrant identity in Germany, Berry et al. (2006) found that 40% of immigrant youth of Turkish descent had a mixed German-Turkish identity with a diffuse (marginalisation) acculturation profile. In the corpus, this is reflected by examples of explicit “neither-nor” auto-stereotyping (10).

- (10) *Bin zu vulgär für alman bubble kann zu gut deutsch für Kanaken bubble und geb zu viel Geld an Freunde für Yahudi bubble*  
(‘I’m too vulgar for the Alman gang, speak German too well for the Kanak gang, and spend too much money on friends for the Jewish gang’) [Alman=snob, Kanak=bad grammar, Jews=avarice]

Obviously, in a multi-cultural setting, not all stereotypes are of equal importance. Thus, personality stereotypes about character and behaviour can present more of a barrier than preference stereotypes about food and music. The most common *Alman*-triggered personality stereotypes encountered in our corpus are listed in table 1.

Table 1: *Alman/Kanak* hetero- and auto-stereotypes

ALMANS (HETERO-STEREOTYPE)	KANAKS (AUTO-STEREOTYPE)
small, nuclear family	large, extended family
laissez-faire upbringing	severe upbringing
free sex	severe morals
sexy females	sexy males
rich & penny-pinching	poor & generous
bureaucratic, theoretical	actions, gut feeling
high-brow talking	slang code
performance-ambitious	knows how to enjoy life
eco-freaks, vegans, NGOs	down-to-earth

7 Fenerbahçe FC, one of the most famous football clubs in Turkey

In addition, there were many examples of what could be called *preference* stereotypes for both Almans [A] and Kanaks [K], e.g. regarding food (A: cheese & jam, mashed potatoes & fish sticks, K: pita bread, melon), cars (A: VW Passat, K: C-Class Merc), hair (A: bad hair, K: stylish hair) or clothing (A: Socks in sandals, check shirts, hiking clothes & rucksack, bicycle helmet, K: Adidas shoes, Alpha jackets, jogging trousers, Nike cap).

### 3.2. Statistics

In order to complement our qualitative stereotype analysis with a quantitative approach, we ran a corpus search for adjectives with either *Alman* or *Kanacke/Kanack/Kanak* as a dependency head. The ranking frequencies in table 2 are a variant of relative frequencies, weighted for actual occurrence. They are calculated as in-sample frequencies squared, divided by "normal" frequency (of the word in a neutral corpus). High values mean that the attribute in question has a strong co-occurrence association with its head and is typical of it<sup>8</sup>. Attributes with a negative sentiment value are in red, positive attributes are in green. Adjectives are indented with a tab if they are more typical for Kanaks, and unindented if they are more typical of Almans. Asterisks mark the difference in ranking frequency by order of magnitude, with a factor 3 difference as a minimum. For instance, *unlustig* ('disagreeable') has two asterisks, because its ranking value for *Alman* is between 10 and 100 times higher than the one for *Kanacke*.

Table 2: Typical adjectives associated with Alman and Kanacke (ranking frequencies, \*/\*\*/\*\* = orders-of-magnitude difference)

& ALMAN & KANAC?KE?		ALMAN (6548)	KANACKE (4972)
<i>groß*</i>	'big'	32797	4045
<i>drecks*</i>	'dirt-' (slur)	1455	5922
<i>weiß**</i>	'white'	3761	25
<i>asozial***</i>	'antisocial'	6	1408
<i>link**</i>	'left-wing'	1347	40
<i>inner**</i>	'internal'	1046	48
<i>unlustig**</i>	'disagreeable'	985	14
<i>hässlich*</i>	'ugly'	269	971
<i>dumm*</i>	'stupid'	132	528
<i>behaart***</i>	'hairy'		467
<i>hübsch**</i>	'pretty'	8	440
<i>letzt*</i>	'last'	431	68
<i>krass</i>	'extreme'	354	414

8 A word with an in-concordance frequency of 1%, and a norm frequency of 1:10,000 will receive a ranking value of 1. The same goes for a 10% word with a 1:100 norm frequency or an 0.1% word with a 1:1,000,000 norm frequency.

<i>schön**</i>	'beautiful'	23	340
<i>dreckig*</i>	'dirty'	76	333
<i>süß*</i>	'cute'	45	317
böse	'evil'	269	116
übel*	'evil'	209	58
<i>schwul*</i>	'gay'	79	194
<i>geil**</i>	'hot'	18	186
<i>fies***</i>	'nasty'	<1	178
fett	'fat'	104	145
<i>fake***</i>	'fake'	<1	116
besoffen*	'drunk'	107	12
nervig*	'irritating'	102	25
blöd*	'stupid'	93	19
echt	'real'	85	41
<i>einzig*</i>	'only'	7	81
<i>eklig</i>	'disgusting'	45	79
<i>fucking***</i>	'fucking'	75	<1
<i>notgeil***</i>	'horny'		68
wütend**	'angry'	59	4
verfickt**	'fucked-up'	59	4
<i>almanisiert***</i>	'almanized'		58
<i>eifersüchtig***</i>	'jealous'		58
<i>richtig**</i>	'real'	52	3
türkisch	'Turkish'	52	32
deutsch	'German'	52	31
<i>muslimisch*</i>	'Muslim'	2	48
<i>gutaussehend*</i>	'good-looking'	<1	48
sogenannt	'so-called'	45	32
<i>stabil</i>	'stable'	14	42
horny	'horny'		40
<i>stinkend*</i>	'stinking'	11	40
<i>männlich</i>	'male'	28	40
scheiß	'shit-' (slur)	39	19
blond	'blond'	37	

As can be seen from the asterisk allocation, the association overweight is generally stronger for *Kanacke* than for *Alman*, with only one 3-star association on the *Alman* side (*fucking*)<sup>9</sup>, indicating that *Alman* is a somewhat more neutral term in the sense of “less stereotyped”. On a sentiment polarity scale, the picture is more nuanced in an interesting way. Thus, *Alman* does have more of the neutral attributes, not least at the top of the list (*big, white, left-wing, inner*), and fewer stars on its negative attributes (*disagreeable, evil, drunk, irritating, stupid, angry*), while it does not have any overweights for positive attributes. It is also interesting that some of the negative *Alman* adjectives are not stereotypes, but dislike markers (*fucking, verfiickt, scheiß*). The attribute list for *Kanacke*, on the other hand, is full of non-neutral stereotypes, both negative and positive. The negative ones top the list, and there are many with 3 stars (*anti-social, hairy, nasty, fake, horny, jealous*). Positive attributes are further down, and have fewer stars, but unlike for *Alman*, they do occur (*pretty, beautiful, cute, hot, good-looking*). Interestingly, all of the latter, and some of the negative adjectives (*horny, jealous, gay, hairy, ugly*) focus on bodily aspects with sexual overtones, while *alman*s are perceived as drab, political, irritating and without sexually likeable or dislikeable attributes.

The corpus is annotated for adjective polarity with Q+/Q- tags. For adjectives with antonyms (*big/small, hot/cold*) this is not, or only metaphorically related to sentiment polarity. For quality adjectives, however, polarity is a safe indicator of sentiment. So we lumped 8 semantic classes of this type (‘affect’, ‘approve’, ‘like’, ‘quality’, ‘skill’, ‘manner’, ‘power’, ‘psyche’) as a sentiment indicator. Table 3 shows how often adjective dependents of the lemmas *Alman*, *Kanacke* and *Deutscher* are positive or negative, respectively.

Table 3: sentiment polarity

	ADJ-MODIFIED(%)	Q+ (%)	Q- (%)	RATIO Q+/Q-
<i>Alman</i>	11.0	9.4	15.0	0.63
<i>Kanacke</i>	29.3	10.5	14.1	0.74
<i>Deutscher</i>	10.3	13.8	13.0	1.06
all nationality or ethnicity nouns	9.0	14.3	8.7	1.64
all +HUM nouns	14.9	16.0	8.2	1.95

The numbers indicate that *Alman*, and even more so *Kanacke*, have a clear overweight of negative attributes (+/- ratio smaller than 1), and are almost twice as likely to occur with negative attributes than other nationality nouns or human nouns in general. Interestingly, *Deutscher* (German), while just about managing a positive +/- ratio, is also more negatively attributed than other nationality nouns. The negativity difference between *Kanacke* and *Alman* is compounded by the fact that the former is almost three times (29.3% vs. 11%) as likely to get an attribute in

<sup>9</sup> The word is an exception, as it is a swear word rather than a stereotype.

the first place. The term's negative overweight is therefore much more pronounced in absolute than in relative terms.

As a complementary approach to sentiment analysis, we used emojis. In the corpus, text emoticons and some relevant emojis<sup>10</sup> are grouped in 9 super-categories, e.g. 'emo-happy' or 'emo-sad', other emojis have individual "lemmas". As a measure of "relatedness", table 4 shows the word vector similarities, in %, between emoji lemmas and a selection of immigrant minority terms. The vector model was a CBOW model (continuous bag of words) trained on content word lemmas (N, ADJ, V, PROP), using the word2vec neural-net method (Mikolov et al. 2013) and the TensorFlow suite (Abadi et al. 2016)<sup>11</sup>. In order to force rare, but relevant words to the surface in the model, and to reduce training time, we used a minority-filtered subcorpus. 100% similarity would mean that two words occur in exactly the same contexts. Obviously, nouns and emojis can't be real synonyms, so here ~40% must be considered a very strong correlation (boldface in the table).

Table 4: Emoji sentiment

	ALMAN	KANACKE	KANAKE	KANAK	DEUTSCHER	ASYLANT	AUSLÄNDER	FLÜCHTLING	MIGRANT	MOSLEM	MUSEL	MOHAMMEDANER	TÜRKE
emo-happy	37	32	31	28	28		28			30	33		30
emo-love	<b>40</b>	37	37	36							32		35
emo-laugh	<b>39</b>	30	29		30					30	<b>40</b>	30	
emo-sad	<b>39</b>	<b>38</b>	36	<b>41</b>							33		
emo-horror	28									31	35	35	
emo-angry						28					<b>40</b>		
emo-surprise	27										29	29	
emo-wink	32	28	28		35		35	29		32	36	28	30
emo-skeptical	34		27		35					30	29		29
thumbs-up										28	32		28
thumbs-down						30			27		39		
oncoming fist	29	33	30			28					34		28
middle finger	28	30	27								<b>40</b>		28

<sup>10</sup> The term *emoji*, in a narrow sense, refers to pictograms, as opposed to character-based text emoticons. In a broader sense, the two are here classified together, into groups of *emoji lemmas*.

<sup>11</sup> dimensions = 100, iterations = 10, vocabulary = 50,000

vomiting						32		27	30		37		
nauseated											39		
goat											39		34
monkey									33	39	31		
ogre						34		28	27	38	33		
pig-face									37	49	34	34	
rat									32	47	35		
camel										37	31		

*Alman* has higher scores for positive emoticons than the *Kanacke* terms, but neither has a strong correlation for “angry”, “horror” or “nausea”, and only a moderate correlation with “oncoming fist” and “middle finger”. Both are moderate terms compared to a real slur like *Musel* (‘Muslim’), which tops the table for most negative emojis. In addition, demeaning animal emojis correlate with all words for muslim, slur or otherwise. Interestingly, wink-emoticons are more prominent for *Alman/Kanacke* than angry/horror emoticons, a finding in line with the commonness of *Alman* jokes (chapter 5).

#### 4. *Alman* hate speech

While most corpus occurrences of *Alman* do not go beyond (negative) stereotyping, irony and mild belittling, there are also many examples of much more demeaning and aggressive usage, as well as outright hate speech<sup>12</sup>, both at the individual (11a-11c) and group levels (12), sometimes even involving religious connotations (13). In their work on Right-wing German HS on Twitter, Jaki & De Smedt (2019) point out that the main concern about online HS is not that it is illegal (it mostly isn’t) but that it is an indicator of society’s polarization (2019: 8). In this light, *Alman*-targeting HS presents an important complementary angle on the other, usually neglected side of this polarization.

(11a) *Brich lieber deine Ausbildung ab sonst verprügeln dich Alman noch die 3 Jährigen Kanackenkinder im Kindergarten du nuttensohn*  
 (‘Better to stop your (pedagogs) education, or else the 3-year-old kanak children will beat you up in Kindergarten’)

<sup>12</sup> One way of extracting HS candidates from the corpus is to search for sentences, where HS keywords, negative emojis or derogatory prefixes co-occur with the terms *Alman*. The Corpuseye interface supports this by allowing chained searches, where the first search is treated as a bag-of-words search (for HS terms) to create an on-the-fly mini-corpus that the second (main) search then operates on. However, to avoid the possible bias involved in pre-selecting HS terms, we also inspected long lists of less constrained *Alman* concordances in order to locate more complex and less lexical examples of HS, like (11b), where the immediate left-hand context of *Alman* is positive (*geil* [‘hot’]), and the negative right-hand context, *khab* [‘hore’], is a Turkish term difficult to come up with beforehand.

- (11b) *Oha geile alman kahb mit Hund gesehen würde aber eher Hund ficken als sie lel*  
 ('Oha seen hot alman hore with dog, but would rather fuck dog than her lel')
- (11c) *mädchen im m dirndl gesichtet hoffentlich kriegst du lungenentzündung scheiss alman*  
 ('seen a girl in a dirndl<sup>13</sup>, I hope you get pneumonia fucking Alman')
- (12) *warum sind fast alle almans hurensöhne wirklich deutschland voll geiles land aber diese almans verkacken alles*  
 ('why are almost all almans sons-of-bitches, really Germany totally cool country but these almans fuck it all up')
- (13) *brudah vallah dise almanpicos denken so puizei kann was machen und so aber nein hahahah di sind halt dum wie schwein hahahahaha von schwein esen zu schwein geworden hahahahaha rihtig dum almanpicos*  
 ('bro these alman sissies think police can do something but no haha they are just stupid like pig haha eat pig become pig haha really stupid alman sissies')

#### 4.1. Othering

Othering (e.g. "what is it with you people?) is a phenomenon commonly found in HS utterances. Typical linguistic surface markers are the use of *ihr/euch* ('you') and *wir/uns* ('we'/'us') in combination with outgroup/ingroup identifiers ('you almans', 'we kanaks'), all of which lent themselves to straight-forward word- and lemma-based corpus searches. Besides outright hate speech (14), othering also occurs in connection with positive vs. negative stereotyping of ingroup and outgroup (15) and in minority counterspeech<sup>14</sup> (16). Without the contrasting *ihr/euch*, *wir/uns* is likely to be auto-stereotyping rather than othering (17).

- (14) *Ihr Almans seid am Ende, ihr Opfas wir machen euch platt. Ihr sterbt aus, weil ihr keine Kinder habt. Kirchen werden zu unseren Moscheen. Ihr seid bald nich mehr da.*  
 ('You almans are finished, you losers we wipe you out. You die out because you don't have children. Churches will be mosques. You'll soon be gone.')
- (15) *wir kanaken haben halt respekt vor ältere personen natürlich gibt özil erdogan die hand ... weil ER der prääsident seiner wurzeln ist | ihr almans beleidigt euren eigenen vater und zeigt euch gegenseitig an | da ist der unterschied*  
 ('We kanaks respect elders, of course Özil shakes hands with Erdogan ... because HE is president of his origins | you almans insult you own father and denounce each other | that is the difference')

<sup>13</sup> close-fitting bodice featuring a low neckline

<sup>14</sup> The example is interesting, because unlike ordinary counter-speech (CS), it is not vented by well-meaning members of the (discriminating) majority, but rather by the minority itself. Also, a kind of counter-attack (counter-othering) is used rather than the normal CS method of pitching positive views against negative ones. Here, the discourse is cross-group, while most CS happens faction-to-faction within the majority, with a risk of inciting even more HS from the scolded faction, as claimed by Ernst et al. (2017) for comments on CS YouTube videos perceived as propaganda.

- (16) *Wenn du behauptest das meine Familie Mafia ist dann hast du richtig gelitten ich bin gerne Mafia wenn ich mit der Familie gleich gestellt bin sowas wie Loyalität, Stolz und Zusammenhalt kennt ihr Almans nicht*  
 ('If you say that my family is mafia you are a real moron, I don't mind being mafia if that is what family is, things like loyalty, pride and solidarity you amans don't know')
- (17) *wir kanaken sind rudeltiere, ich muss leute rufen*  
 ('we kanaks are flock animals, I need to get my friends')

#### 4.2. Generalizations

Though not proof by themselves, ethnic generalizations are symptomatic of HS. Typical grammatical vehicles are indefinite plural<sup>15</sup> (18), demonstrative plural (19a-19c), the 'all' quantifier (19b) or a definite article without attributes (19c).

- (18) *almans bleiben halt immer almans, der eine jagt kanaken durch straßen, der Andere macht's virtuell*  
 (,Almans are almans, one chases kanaks in the streets, another does it online')
- (19a) *lak immer diese almans dümmste mensch auf Erde*  
 ('bro always these almans, stupidest people on earth')
- (19b) *Wallah diese almanns haben doch schaden*  
 ('Bro these almans are all nuts')
- (19c) *Würde eig jetzt gerne dann diese almans abfucken bis die sich löschen oder ich fahre dahin und suche mir einzel jeden davon und fuck die real ab*  
 ('would like to fuck these almans until they delete themselves, or I go there find each of them and fuck them for real')
- (20) *Ich sage nur Mashallah und alle Almans sind dumm.*  
 ('I just say mashallah and all Almans are stupid')
- (21) *Der Alman ist und bleibt haram.*  
 ('Almans will always be haram/sin')

A corpus search for n-grams starting with the generalizing "Almans sind" ['Almans are ...'] found the following hateful examples, ordered by frequency. The derogative elements feature the usual dehumanizing metaphors (trash, dirt, animals), faecal terms and demeaning adjectives (evil, mangy, mentally ill). In addition, and fairly high up in the ranking, there are the (for Germans) particularly sensitive attributes *nazi*, *racist* and *fascist*. For the tweeter, these are probably target-specific.

<sup>15</sup> Examples of this construction can be found in the corpus by either specifying a word field left of the demonym plural syntactically (as "not prenominal function") or morphologically (as not article, adjective or determiner), or by exploiting the fact that the corpus annotation inserted a special dummy field for "beginning of sentence". A higher density of "good" examples can be achieved by constraining the demonym search field with the syntactic function of 'subject'.

Almans sind

*Abfall/abfall/abfall für mich* ['scum'/'trash']  
*so unfassbarer dreck* [,unimaginabel dirt']  
*so geisteskrank* [,so mentally ill']  
*nazis/nadsis/nazienkel/alles nazi* [,nazis']  
*echt voll rassistisch* [,racist']  
*rassisten (und gehören)* [,racist and should be ...']  
*24/7 rassistisch* [,24/7 racist']  
*faschos/alle antisemiten* [,fascists']  
*doch keine menschen* [,not human']  
*einfach nur scheiße* [,just shit']  
*seelenlose kreaturen* ['soal-less beasts']  
*die bösen* ['the evil ones']  
*so geier* ['vultures']  
*einfach unnormale cucks* ['abnormal "not manly ones"']  
*ekelhaft/scheise* ['nauseating'/'shit']  
*solche hurensöhne* ['sons-of-prostitute']  
*nur ein räudiges* [,just mangy ...']  
*übelst lauch* ['most evil leeks']  
*die größten heuchler/doppelmoralisch* [,hypocrits']

## Alman jokes

Humour, of one kind or another, is present in a large portion of *Alman* corpus hits. Lexically, it can take the form of word creations (see ch. 2 and (22)), word allusions (23: PoC, WoC), irony (24) or connotation clashes (23, about Merkel: Mummy [good] <-> nazi [bad]). Structurally, contrasting of Almans and Kanaks is very common (25)<sup>16</sup>.

- (22) *#Kopftuchdebatte ist **Almanitis** im Endstadium.*  
 ('#headscarf-debate is final-stage almanitis')
- (23) *Ich als marginalisierter **Alman of Color** sage: Alles Nazis, auch Mutti.* 😏  
 ('As a marginalized Alman-of-Color I say: They are all nazis, Mummy, too 😏')
- (24) *nur almans kennen die mechanismen von mobbing, wir ausländer wachsen in raumkapseln auf<3*  
 ('only Almans know how to mob, we foreigners grow up in outer space <3')
- (25) *almans sehen nach friseurbesuch aus wie kanaken vor*  
 ('after going to the hairdresser's, almans look like kanaks before going there')

<sup>16</sup> To identify candidates for contrasting jokes in the corpus, the aforementioned technique of chained searches is quite effective. Here, a *Kanack* search is performed on top of an *Alman* search or vice-versa. In the most common joke scheme, both terms are in the plural and subjects, allowing for further, tag-based search constraints.

The contrasting technique often follows a more or less formalized joke scheme: *Almans are/do/like X while kanaks are/do/like Y*. The humorous effect is achieved by black-and-white exaggeration, but part of the jokes' appeal is that they are anchored in real-life examples and common preconceptions. As a corpus, the *Alman-Kanak* jokes provide interesting insights in both hetero- and auto- stereotyping at the same time:

- (26) *Kanaken immer halbe Klasse miternährt während Almans mit Geodreieck und Parabel Brötchen genau ausmessen damit keiner zu viel bekommt*  
 ('Kanaks always feed half the class, while Almans share rolls with a ruler so nobody gets too much')

In many cases, Alman jokes come with an initial topic opener. These are either simple NPs (27-28) or *wenn/wie*-clauses (29-30).

- (27) *Erfolg bei Almans: 1,0 Abi | Erfolg bei Kanacken: Meros Lieder 10h vorher und überall damit rum pushen*  
 ('success for almans: all-A grades | success for kanaks: Mero's song 10 hrs in advance and show them off')
- (28) *almans zum einzellkind: möchtest du heut abend lasagne oder schnitzel? | kanaken zum 12. sohn: willst du jetzt essen oder schläge?*  
 ('Almans to single child: would you like lasagne or schnitzel for dinner? | kanaks to 12th son: will you eat or have a spanking?')
- (29) *wenn Besuch übernachtet | almans: schön eingerichtetes Gästezimmer, morgens frische Brötchen, Handtücher und frische Wäsche zum Duschen | kanacks: kinder werden aus Zimmer gejagt, morgens Pide von gestern getoastet, fetter Onkel schwitzt Bett voll und verstopft Klo*  
 ('when guest stay overnight | almans: nice guest room, newly-baked rolls for breakfast and fresh towels for the shower | kanaks: children (pl.!) thrown out of their room, yesterday's pide (bread) toasted for breakfast, fat uncle sweat-marks the bed and shits the toilet blocked')
- (30) *wie almans kanakenwörter aussprechen: - schüsch - maschaller - waller | wie kanaken almanwörter aussprechen: Rindfleischetikettierungsüber.....überw.....üüüüübeeeeerrwaaaaa junge vallah was das für scheiße amk diese wort ist halbe buch wer kann das lesen amk*  
 ('how almans pronounce kanak words: shüsh - mashaller - waller | how kanaks pronounce alman words: beef-labeling-monii...moniitooor... boy vallah what that shit amk (*amina koyim* - Turkish: 'in-your-cunt') these word is half book who can read that amk')

As can be seen, the implied stereotypes are quite diverse. Thus, (26) targets German avarice and Kanak generosity, as well as the idea of German being meticulous in an exaggerated manner, while (27) contrasts the German focus on performance with immigrants focus on social standing in your peer group. (28) and (29) take an ironic view on differences in wealth, family patterns and children's' status at the same time, not to forget food stereotypes in (28). (30), finally, addresses the afore-mentioned prejudice against Kiez-Deutsch as an inferior language variant simply caused by not knowing German properly. Another common stereotype is German fear of dirt and infection and the belief that immigrants are a source of both (31).

- (31) *Wenn Kanaken Obst essen - Kaufen es und zerfetzen es direkt | Wenn Almans Obst essen - Karl Gerhard Theodor Gustav Liebling wasch bitte den Apfel nochmal vorher ab man weiß ja nie welcher drekiger Kümmel Türke den vorher mit drekigen Händen berührt hat. Danke*  
 ('When kanaks eat fruit - buy it and lacerate it at once | when almans eat fruit - K.G.Th.G. darling wash the apple once more, you never no which dirty "cumin Turk" has touch it with his dirty hands')

The two holiday jokes (32-33) are quite different. The first paints a (positive) self-image of kanaks being relaxed and blessed with a down-to-earth common sense and disregard for petty rules. The second joke is a self-ironic description of how immigrant "holidays" may mean helping your left-behind family with labor and coveted food.

- (32) *Alman im Urlaub: | 1. um 6 Uhr aufstehen | 2. Handtuch auf Liege platzieren | 3. spazieren gehen | 4. Handtuch suchen | Kanacken: | 1. um 15 Uhr aufstehen | 2. schöne Schattenplatz suchen | 3. Handtuch von Alman wegschmeißen und sich auf Liege chillen*  
 ('Alman on holiday: 1. get up at 6 o'clock | 2. place towel on sunbed | 3. go for a walk | 4. search towel | Kanaks: 1. get up at 3 p.m. | 2. find a nice shady place | 3. throw away the Alman towel and chill on sunbed')
- (33) *Urlaub Almans: - teure Skiurlaub - jede Ferien weil reich wie Dagobert Duck - Mallorca zweite Zuhause | Ausländer: 6 Wochen in unterentwickelte Dorf Heu wie Esel schleppen -20 kg Nussbeisser Gepäck -in Alanya aufgeile Russinnen geiern*  
 ('Alman holiday - expensive ski holidays all the time because rich like Dagobert Duck - second home on Mallorca | Foreigner: carry hay like donkey for 6 weeks in underdeveloped village - 20 kg Nussbeisser luggage - lascivate about hot Russian chicks in Alanya')

Not always the opposition pair is *Alman* vs. *Kanack* - sometimes an Alman joke exploits other contrasts, like the political CS joke in (34).

- (34) *Linksextreme zünden Autos an, Rechtsextreme Ausländer. Almans finden Linke schlimmer weil sie Autos besitzen aber keine Ausländer*  
 ('Left-wing extremists burn cars, right-wing extremists burn foreigners. Almans think left-wingers are worse, because Almans own cars, but no foreigners')

Finally, in-group behaviour can be contrasted with inter-group situations (35):

- (35) *kanacke zum kanacken: k1: ich hab deine schwester gefickt - k2: du hurensohn ich bring dich um | kanacke zum alman: k1: ich hab dein schwester gefickt - almann: habt ihr auch verhütet ? 😊*  
 ('kanak to kanak: k1: I fucked your sister - k2: you son-of-prostitute, I kill you | kanak to alman: k1: I fucked your sister - alman: have you used contraception?')

## 6. Conclusion

As a corpus-linguistic tool, the Kiezdeutsch term *Alman* provides a rare opportunity to examine minority ("counter-") hate speech in German rather than e.g. Arabic or Turkish. Thus, in its morphosyntactically and semantically annotated form, our social-media corpus allowed us to identify derogatory language and outright HS targeting the German host community (involving e.g. the concepts of *Nazi* and sexual honor) as well as [negative] stereotypes linked to it (e.g. over-zealous, bureaucratic, eco-crazy, penny-pinching). In addition, we were also able to verify the use of othering and generalization strategies as a means of verbal aggression against the *Alman* out-group. However, while generally used with negative connotations, the word *Alman* does not appear to be as negative as the equally stereotyped counterpart *Kana(c)ke* or the religious-ethnic slur *Musel* ('Muslim'), judging from a vector space analysis involving negative and animal emojis. On the other hand, *Alman* is rarely used for auto-stereotyping and thus lacks the positive stereotypes sometimes associated with *Kanak* (e.g. sexiness, virility) when used as an in-group term.

One indication of extensive stereotyping in the corpus is the fact that linguistic negative attributes (adjective dependents) of both *Alman* and *Kana(c)ke* are often specific (e.g. about appearance or character) rather than general (e.g. *scheiß, drecks*), and quite different for the two groups. Notwithstanding negative stereotyping, however, humorous contexts are much more common than outright demeaning ones, to the point of manifesting in a formalized joke scheme, where traits and behavior of *Almans* and *Kanaks* are compared. In our corpus study, jokes were not only a good source of hetero-stereotypes, but also laden with (exaggerated) auto-stereotypes and a healthy portion of self-irony.

## References

- Abadi, Martin et al. 2016. TensorFlow: A system for large-scale machine learning. In *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*. 265–283.
- Baumgarten, Nicole; Eckhard Bick; Klaus Geyer; Ditte Iversen; Andrea Kleene; Anne Vibeke Lindø; Jana Neitsch; Oliver Niebuhr; Rasmus Nielsen; and Esben N. Petersen. 2019. Towards balance and boundaries in public discourse: expressing and perceiving online hate speech (XPEROHS). *RASK - International Journal of Language and Communication* 50. 87–108
- Bell, Jeannine. 2009. Restraining the heartless: racist speech and minority rights. *Indiana Law Journal* 84. 964–980.
- Berry, John W.; Jean S. Phinney; and David L. Sam. 2006. Immigrant youth: acculturation, identity, and adaptation. *Applied Psychology* 55(3). 303–332
- Bick, Eckhard. 2020. An annotated social media corpus for German. In Nicoletta Calzolari et al. (eds.) *Proceedings of the 12th International Conference on Language Resources and Evaluation, LREC 2020* (Marseille, May 2020). ACCL/ELRA. 6129–6137
- Ernst, Julian; Josephine B. Schmitt; Diana Rieger; Ann-Kristin Beier; Peter Vorderer; Gary Bente; and Hans-Joachim Roth. 2017. Hate beneath the counter speech? A qualitative content analysis of user comments on YouTube related to counter speech videos. *Journal for Deradicalization* 10. 1–49.
- Foxman, Abraham, and Christopher Wolf. 2013. *Viral hate: containing its spread on the Internet*. New York: St. Martin's Press.
- Freywald, Ulrike; Katharina Mayr; Tina Özçelik; and Heike Wiese. 2011. Kiezdeutsch as a multiethnolect. In Friederike Kern & Margret Selting (eds.) *Ethnic styles of speaking in European metropolitan areas*. Amsterdam: John Benjamins. 45–73
- Jaki, Sylvia, and Tom De Smedt. 2019. Right-wing German hate speech on Twitter: analysis and automatic detection. arXiv preprint arXiv:1910.07518.
- Martiny, Sarah E.; Sog Yee Mok; Kay Deaux; and Laura Froehlich. 2014. Effects of activating negative stereotypes about Turkish-origin students on performance and identity management in German high schools. *Revue internationale de psychologie sociale* 27(3). 205–225. [https://www.cairn.info/revue-internationale-de-psychologie-sociale-2014-3-page-205.htm]
- Mandel, Ruth. 2008. *Cosmopolitan Anxieties: Turkish challenges to citizenship and belonging in Germany*. Durham: Duke University Press.
- Mikolov, Tomas; Ilya Sutskever; Kai Chen; G. S. Corrado; and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In C.J.C. Burges et al. (Eds.) *Advances in neural information processing systems - Proceedings of NIPS 2013*. 3111–3119.
- Wiese, Heike. 2011. Führt Mehrsprachigkeit zum Sprachverfall? Populäre Mythen vom „gebrochenen Deutsch“ bis zur „doppelten Halbsprachigkeit“ türkischstämmiger Jugendlicher in Deutschland. In Ozil, Şeyda; Hofmann, Michael; and Dayioğlu-Yücel, Yasemin (eds.) *Türkisch-deutscher Kulturkontakt und Kulturtransfer. Kontroversen und Lernprozesse*. Göttingen: V&R Unipress. 73–84.