

This **Discussion Paper** is an open peer review version that we do not recommend to cite. Submissions that have passed the peer review process are published as full articles on [www.jfml.org](http://www.jfml.org).

the editors

**jfml**

Journal für Medienlinguistik  
Journal for Media Linguistics

Diskussionspapier 6 (2025):  
eingereicht am 19.08.2025

Gutachten und Kommentare unter:  
<https://dp.jfml.org/2025/opr-flausch/>

Discussion Paper

1 **Flausch**

2 **Unterstützende Sprachhandlungen in sozialen Medien**

3 *Yulia Clausen & Tatjana Scheffler*

4 **1 Einleitung**

5 Soziale Medien stellen einen festen Bestandteil des Alltags  
6 dar. Ihre weitreichende gesellschaftliche Relevanz lässt sich  
7 unter anderem darauf zurückführen, dass sie den Zugang zu  
8 virtuellen Lebenswelten eröffnen (Rieger/Schäfer/Tuschling  
9 2021). In solchen Kontexten finden Nutzer\*innen häufig die  
10 Freiheit, sich neu bzw. anders zu gestalten (Kersten/Lotze  
11 2019) und dadurch unterschiedliche (virtuelle)  
12 Persönlichkeiten auszuleben (Burnett 2017; 2019). So können  
13 sie an verschiedenen virtuellen Gemeinschaften (oder  
14 Communities; Rheingold 1994) teilhaben bzw. neue gründen  
15 und sich auf diese Weise vielfältige virtuelle Lebenswelten  
16 erschließen. Diese stehen jedoch nicht im Widerspruch zu  
17 physischen Welten, sondern erweitern sie, insbesondere  
18 durch die Pluralität der Kontexte und die sozialen  
19 Gemeinschaften, in denen sie verortet sind. Virtuelle  
20 Gemeinschaften verstehen wir als solche, die sich unabhängig  
21 von örtlichen und zeitlichen Beschränkungen durch  
22 Gruppenbildungsprozesse und soziale Interaktion  
23 herausbilden, zum Beispiel in Form von Interessensgruppen  
24 oder Fan-Communities.

25 Virtuelle Phänomene sind eng mit dem Alltagsleben  
26 verflochten, sodass sich virtuelle und nicht-virtuelle  
27 Gemeinschaften sowie Interaktionen zunehmend schwer  
28 voneinander abgrenzen lassen. Das hat zur Folge, dass ein  
29 immer größerer Anteil sprachlicher Kommunikation digital  
30 mediiert und nicht von Angesicht zu Angesicht im physischen  
31 Raum stattfindet. Wir gehen von der Beobachtung aus, dass

## Clausen & Scheffler: Flausch

Individuen täglich viel Zeit in sozialen Medien verbringen<sup>1</sup> und große Mengen sprachlicher und nicht-sprachlicher Interaktionen im Internet stattfinden. Digital vermittelte sprachliche Kommunikation gilt oft als weniger stark sozial eingebettet im Vergleich zu den nicht-digital mediierten Kommunikationsformen. Dies wird häufig als Erklärung für eine gewisse Enthemmung in sozialen Medien und vielverbreitete Hassrede angeführt (Suler 2004; Jaki 2023). Allein reicht diese Perspektive jedoch nicht aus, um die hohe Popularität sozialer Medien zu erklären: Wären diese tatsächlich reine Orte sozialer Vereinzelung, an denen Nutzer\*innen sich Hass und Hetze aussetzen, wäre es fraglich, warum Individuen dort regelmäßig Zeit verbringen und Inhalte sowie ihre Lebenswelten teilen. Wir vertreten die Hypothese, dass Hassrede nur einen Teil virtueller Kommunikation ausmacht, und virtuelle Communities häufig von positivem Austausch und gegenseitiger Unterstützung geprägt sind.

Wir bezeichnen diese Form des unterstützenden, gruppenkonstruierenden Austausches als „Flaschrede“ (engl. *candy speech*). Wie unsere Korpusdaten zeigen, handelt es sich dabei um einen zentralen und allgegenwärtigen Aspekt virtueller Kommunikation (s. Abschnitt 4). Obgleich Flaschrede in sozialen Medien mindestens ebenso stark präsent ist wie Hassrede, erfährt sie im Gegensatz zu letzterer deutlich geringere Wahrnehmung und Aufmerksamkeit. Zu Hassrede und ihrer (automatischen) Erkennung liegt eine Vielzahl an Studien vor (s. u. a. Hosseinmardi et al. 2015; Brodnig 2016; Schmidt/Wiegand 2017; Wiegand et al. 2019; Ashraf et al. 2021). Positive, unterstützende Onlinesprache ist hingegen in der (Computer-)Linguistik noch weitgehend unerforscht (für Ausnahmen siehe Abschnitt 3.2). Ziel der vorliegenden Studie ist es, dieses Ungleichgewicht zu adressieren.

Im Folgenden erarbeiten wir, basierend auf der Etymologie des Begriffs „Flasch“, eine Definition dieses Phänomens sowie seiner Untertypen. Wir leiten Erkennungsmerkmale ab und untersuchen das Vorkommen sowie die Verwendung von

---

<sup>1</sup> Im Schnitt verbringen Nutzer\*innen weltweit täglich 2 Stunden und 26 Minuten auf Social-Media-Plattformen (<https://www.statista.com/statistics/270229/usage-duration-of-social-networks-by-country/>; Stand 2023; abgerufen am 22.05.2025).

## Clausen & Scheffler: Flausch

70 Flauschrede exemplarisch anhand eines Korpus deutscher  
71 YouTube-Sprache (Cotgrove 2023). Abschließend  
72 vergleichen wir Flausch- und Hassrede im Hinblick auf ihre  
73 jeweiligen Klassifikationsmerkmale. Unser Ziel ist es, die  
74 Relevanz von Flauschrede als virtuellem sozialen Phänomen  
75 hervorzuheben und ein Bewusstsein für ihre Präsenz in  
76 sozialen Medien zu schaffen. Die automatische Erkennung  
77 und Sichtbarmachung von Flauschrede kann dabei eine  
78 ergänzende Strategie zur Eindämmung von Hassrede  
79 darstellen und zur positiven Gestaltung virtueller  
80 Kommunikationsräume beitragen.

## 81 **2 Zur Entstehung des Begriffs „Flausch“**

### 82 2.1 Etymologie

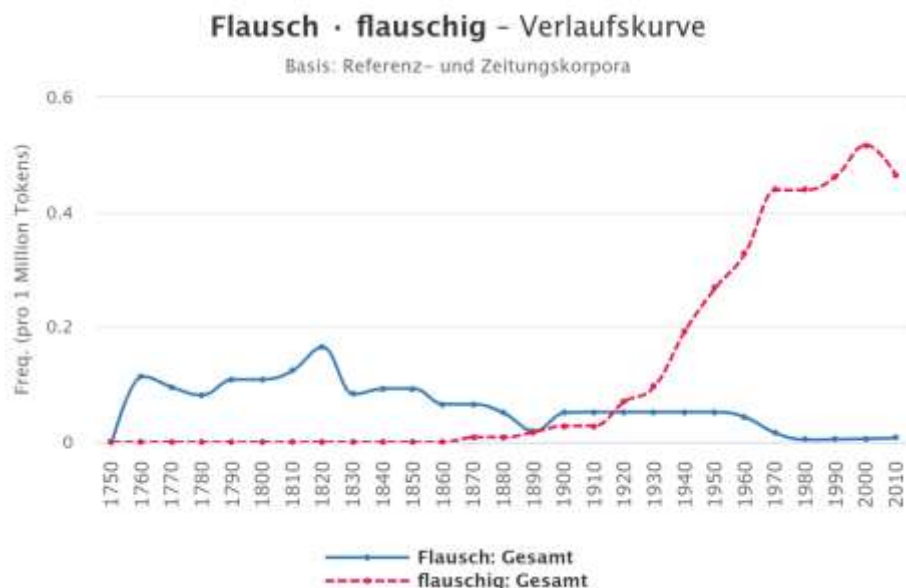
83 Der Begriff „Flausch“ stammt aus dem 18. Jahrhundert und  
84 bezeichnete ursprünglich einen „weichen, haarigen  
85 Wollstoff“.<sup>2</sup> Das Wort leitet sich von *vlūs*, *vlüşch* („wolliges  
86 Schaffell“) ab und bezieht sich auf ein dickes, weiches, stark  
87 gerautes Wollgewebe, das für Mäntel verwendet wurde.  
88 Parallel dazu wird das Adjektiv „flauschig“ („weich, wollig“;  
89 19. Jahrhundert) im wörtlichen Sinne für weiche, kuschelige  
90 Stoffe und Plüschtiere sowie auch für kleine Tiere  
91 verwendet.<sup>3</sup> Das Substantiv ist heute im ursprünglichen Sinn  
92 eher ungebräuchlich und leitet sich in seiner Bedeutung vom  
93 Adjektiv ab (s. Abbildung 1).

---

2 „Flausch“, bereitgestellt durch das Digitale Wörterbuch der deutschen Sprache, <https://www.dwds.de/wb/Flausch> (abgerufen am 12.03.2025).

3 „flauschig“, in: Wolfgang Pfeifer et al., Etymologisches Wörterbuch des Deutschen (1993), <https://www.dwds.de/wb/etymwb/flauschig> (abgerufen am 12.03.2025).

## Clausen & Scheffler: Flausch



**Abbildung 1:** DWDS-Wortverlaufskurve für „Flausch · flauschig“, erstellt durch das Digitale Wörterbuch der deutschen Sprache, <https://www.dwds.de/r/plot/?view=1&corpus=public&q1=Flausch&q2=flauschig> (abgerufen am 07.02.2024).

### 2.2 Flausch in digitalen Medien

In digitalen Medien taucht der Begriff „Flausch“ erstmals im Jahr 2009 im Blog von Stefan Niggemeier in der Rubrik „Flausch am Sonntag“ auf. Beiträge, die 2008 zunächst als „Katzencontent“ veröffentlicht wurden, haben sich im Laufe der Zeit zum sogenannten „Flauschcontent“ entwickelt. Dabei handelt es sich um Video- und Fotobeiträge, die (im wörtlichen Sinne) flauschige Tiere zeigen, seien diese real, gezeichnet oder aus Plüsch.

Spätestens im Jahr 2012 wurde der Begriff von der Piratenpartei aufgegriffen. In der Zwischenzeit hatte er eine deutliche Bedeutungserweiterung erfahren und wurde nicht länger ausschließlich im wörtlichen Sinne als Bezeichnung für Tiere mit weichem Fell verwendet, sondern auch metaphorisch auf Inhalte übertragen, die allgemein als süß, positiv oder emotional wohltuend empfunden werden. In der Terminologie der Piratenpartei wurde Flausch im Sinne eines „erträglichen und netten Miteinanders gemäß der Internet-Netiquette“ definiert und mit Konzepten wie Fürsorge, Entspannung und „MITeinander statt GEGENEinander“ assoziiert. In diesem Kontext gründete die Partei die sogenannte „AG Flausch“, deren Ziel es ist, „Mittel und Wege [zu] entwickeln um für einen flauschigeren Umgang unter den

## Clausen & Scheffler: Flausch

Mitgliedern der Piratenpartei zu sorgen“.<sup>4</sup> In Anlehnung an diese Idee bietet die AG eine „Portion Flausch“ an, sollte es mal „Not am Manne oder an der Frau“ sein.

Damit sind die Begriffe „Flausch“ und „flauschig“ im allgemeinen Sprachgebrauch der Internetgemeinde angekommen und bezeichnen generell unterstützende und positive Beiträge in sozialen Medien und im Netz. Beispielfhaft kann hierfür ein Beleg aus dem NottDeuYTSch-Korpus aus dem Jahr 2017 herangezogen werden (1).<sup>5</sup>

- (1) Geiles video #Glocke aktiv #daumen nach oben 👍  
#hdl ❤️ Fände es flauschig wenn du mich grüßen  
könntest 💋 ❤️

### 3 Definition von Flausch und Flauschtypen

In diesem Abschnitt erarbeiten wir eine Definition von Flausch, insbesondere in Abgrenzung zu verwandten Phänomenen, und beschreiben dessen verschiedene Typen.

#### 3.1 Definition

Wir definieren Flausch als Ausdruck positiver Einstellungen und Gefühle in sozialen Medien gegenüber Einzelpersonen, d. h. Contentcreator\*innen (in unseren Daten Autor\*innen von Videos) und Kommentator\*innen oder deren Beiträgen (Videos und Kommentare). Der Zweck von Flausch besteht darin, andere Personen zu ermutigen, aufzumuntern, zu unterstützen oder zu bestärken. Demnach zielt Flausch darauf ab, das Selbstbild der Zielperson oder -gruppe auf positive Weise zu beeinflussen.

Analog zur toxischen Sprache und zur Hassrede, die häufig durch ihre Effekte (z. B. Einschüchterung, Ausgrenzung oder emotionale Verletzung) sowie durch die Intentionen ihrer Verwender\*innen (z. B. Provokation oder Abwertung) charakterisiert werden (Jaki 2023), betrachten wir den intendierten perlokutiven Effekt des Flauschs, nämlich die Bestärkung des positiven Selbstbildes einer Person oder Gruppe, als definierendes Kriterium dieses Begriffs. Flausch lässt sich demnach als kommunikative Handlung verstehen,

<sup>4</sup> [https://wiki.piratenpartei.de/AG\\_Flausch](https://wiki.piratenpartei.de/AG_Flausch) (abgerufen am 12.03.2025).

<sup>5</sup> Der Beleg stammt aus Datei NDY/202.

## Clausen & Scheffler: Flausch

- 157 die unterstützenden *Face Work* im Sinne Erving Goffmans  
158 leistet. Face Work bezeichnet (sprachliche) Handlungen, die  
159 das soziale Selbstbild – das sogenannte *Face* (Goffman 1967)  
160 – von Sprecher\*innen oder Adressat\*innen positiv oder  
161 negativ beeinflussen.
- 162 3.2 Abgrenzung von ähnlichen Konzepten
- 163 Abschnitt 3.1 zeigte, dass Flauschrede der positiven  
164 phatischen Kommunikation zuzuordnen ist. Als phatisch wird  
165 jene Form der Kommunikation bezeichnet, die der  
166 Herstellung und Aufrechterhaltung sozialer Beziehungen  
167 dient, ohne dabei auf den Austausch inhaltlich relevanter  
168 Informationen abzielen (Klüwer 2011). Im  
169 englischsprachigen Raum existieren mehrere Begriffe, die  
170 unterschiedliche Aspekte phatischer Kommunikation  
171 abdecken und deren Bedeutung sich teilweise mit der von  
172 Flausch überschneidet. Hinsichtlich der Bedeutung und des  
173 Gebrauchs steht der Begriff „Candystorm“ dem Flausch am  
174 nächsten:
- 175 Mit dem Candystorm geht eine Welle des Zuspruchs im  
176 virtuellen Raum einher, z.B. in sozialen Netzwerken,  
177 Microblogs und Blogs sowie Kommentarbereichen von  
178 Onlinezeitungen und -zeitschriften. Er wird evoziert durch  
179 den Moralismus der Informationsgesellschaft und die  
180 Empathie und Euphorie der Netzbürgerinnen und -bürger.  
181 Personen oder Organisationen werden mit Worten des  
182 Zuspruchs und Begriffen wie „Flausch“ bedacht. Das  
183 Gegenteil ist der Shitstorm (Bendel 2019: 37).
- 184 Wie aus dieser Definition hervorgeht, bezeichnet  
185 Candystorm – ebenso wie sein Gegenpol Shitstorm –  
186 wellenartige Ausbrüche entsprechender sprachlicher  
187 Dynamiken. Der Begriff „Flauschrede“ (sowie sein  
188 Gegenstück „Hassrede“) ist dagegen neutraler gefasst und  
189 lässt sich auch auf kleinere Einheiten unterstützender  
190 Kommunikation anwenden, beispielsweise auf einzelne  
191 Kommentare.
- 192 *Empowering language* (bestärkende Sprache) bezeichnet  
193 eine weitere Form phatischer Onlinesprache, deren Ziel es  
194 ist, „die Rechte, die Wahlmöglichkeiten, die  
195 Selbstverwirklichung oder das Selbstwertgefühl von



## Clausen & Scheffler: Flausch

Adressat\*innen zu fördern“ (Njoo et al. 2023: 9335; unsere Übersetzung). Ähnlich wie beim Flausch handelt es sich dabei um unterstützende Sprachhandlungen, die das positive Selbstbild der Adressat\*innen stärken. Da der Fokus hierbei jedoch ausschließlich auf der Stärkung der gefühlten Handlungsmacht der Adressat\*innen liegt, ergibt sich nur eine partielle Überschneidung mit unserer Definition von Flausch. So werden in der zitierten Studie Kommentare wie „And know that it’s ok to be sad. It’s ok to mourn the loss of someone you love - whether they were great or they were awful. Its normal and you shouldn’t feel weird or bad because of it. Good luck :)“ als bestärkend klassifiziert (Njoo et al. 2023: 9334). Obwohl solche Beiträge darauf abzielen, bei den Adressat\*innen ein positives Gefühl hervorzurufen, richten sie sich nur indirekt auf die Unterstützung ihres sozialen Selbstbildes (Face). Viele der in der Studie als *empowering* klassifizierten Aussagen würden daher nicht notwendigerweise unter Flausch im Sinne unserer Definition fallen.

Ein weiterer verwandter Begriff aus der Literatur ist „Smalltalk“ (oder *social talk*). Dieser Interaktionstyp fällt ebenfalls in den Bereich der phatischen Kommunikation und dient weniger dem Informationsaustausch als der „Herstellung von sozialen Kontakten“ (Klüwer 2011: 14). Klüwer (2011) schlägt eine Taxonomie von Dialogakten vor, die auf kooperative soziale Gespräche (Smalltalk) in virtuellen Welten (nach Art von *Second Life*) ausgelegt ist. Diese Taxonomie basiert ebenso auf Goffmans Theorie des Face Work und unterstützenden Dialogakten. Sowohl Smalltalk als auch Flausch haben zum Ziel, Beziehungen zu pflegen und somit das „Gesicht“ des Gegenübers zu wahren. Sie unterscheiden sich jedoch insofern, als Flausch in der Regel deutlich stärkere emotionale Unterstützung bietet (s. Abschnitt 3.1). Smalltalk beinhaltet hingegen beispielsweise sprachliche Akte, die Interesse am Gesprächspartner signalisieren, indem gezielte Nachfragen gestellt werden, ohne dabei eine direkte emotionale Unterstützung zu leisten. Solche Beiträge würden nicht unter Flauschrede fallen, da ein wesentlicher perlokutiver Effekt in der inhaltlichen Auseinandersetzung liegt.

Für die digital medierte Kommunikation wurde zudem der Begriff *hope speech* (Hoffnungsrede) geprägt. Er bezeichnet

## Clausen & Scheffler: Flausch

eine Sprachform, die „positiv, ermutigend, beruhigend, integrativ und unterstützend ist und Menschen inspiriert sowie Optimismus vermittelt“ (Chakravarthi et al. 2022: 379; unsere Übersetzung). Diese Definition zeigt, dass *hope speech* in bestimmten Aspekten mit dem Konzept des Flauschs übereinstimmt, jedoch einen inhaltlich abweichenden Schwerpunkt setzt. Sie bezieht sich ausschließlich auf sprachliche Äußerungen, die der Entschärfung von Feindseligkeiten dienen (Chakravarthi 2020; Palakodety et al. 2020; Garcia-Baena et al. 2023). Im Gegensatz dazu setzt Flausch nicht zwingend eine vorangehende feindliche oder negative Kommunikation voraus, sondern kann auch unabhängig davon in unterstützender Absicht eingesetzt werden.

### 3.3 Typen von Flausch

Für die feingliedrige Klassifikation von Flauschausdrücken greifen wir auf die Taxonomie der Smalltalk-Dialogakte nach Klüwer (2011) zurück. Klüwer berücksichtigt in ihrer Analyse sowohl Formen des Face Work, die auf das eigene Gesicht der Sprecher\*innen abzielen, als auch solche, die auf das Gesicht der Adressat\*innen gerichtet sind. Im Fall von Flausch ist jedoch vor allem das adressatenbezogene Face Work relevant.

Vor diesem Hintergrund übernehmen wir aus Klüwers Taxonomie jene Kategorien, die mit unserer Definition von Flausch übereinstimmen, wie zum Beispiel Komplimente und positives Feedback, und erweitern sie um zusätzliche Typen, die auf der Grundlage unserer Korpusanalyse identifiziert wurden. Im Gegensatz zu Klüwer annotieren wir in der empirischen Anwendung keine gesamten Dialogakte, sondern ausschließlich die darin enthaltenen Flauschausdrücke. Unsere Klassifikation umfasst zehn Kategorien: *affection declaration*, *compliment*, *encouragement*, *gratitude*, *group membership*, *positive feedback*, *sympathy*, *implicit*, *agreement* und *ambiguous*. Die letzten drei stellen keine eigenständigen Flauschtypen dar, sondern decken zusätzliche Fälle ab. Im Folgenden beschreiben wir die einzelnen Kategorien im Detail.<sup>6</sup>

<sup>6</sup> Alle Beispiele entstammen dem in Abschnitt 4 vorgestellten Korpus. Die Kommentare wurden originalgetreu inklusive aller Rechtschreibfehler



## Clausen & Scheffler: Flausch

276 *Affection declaration* – Bewunderungs-, Liebes- und  
277 Zuneigungsbekundungen gegenüber Personen, d. h.  
278 Contentcreator\*innen oder Kommentator\*innen (und nicht  
279 gegenüber deren Beiträgen, wie Videos oder Kommentaren).  
280 Diese Äußerungen müssen nicht notwendigerweise  
281 romantisch gemeint sein.

282 (1) ich liebe dich so sehr würde mich sehr freuen wen du  
283 mir antworten könntest und mich grüßenn könntest  
284 #beste youtuberin  
285 (NDY/202)

286 *Compliment* – Komplimente und freundliche Bemerkungen,  
287 die sich auf andere Personen und nicht auf deren Beiträge  
288 beziehen. Komplimente können Fähigkeiten (z. B. Singen,  
289 Kochen), persönliche Eigenschaften (z. B. Stimme, Aussehen)  
290 oder auf Leistungen (allgemein oder im Kontext eines  
291 bestimmten Videoauftritts) sowie deren Bewertung betreffen.

292 (2) deine stimme ist sehr angenehm . du bist  
293 wunderschön !  
294 (NDY/020)

295 *Encouragement* – Äußerungen, die darauf abzielen, andere  
296 Personen (Contentcreator\*innen und Kommentator\*innen)  
297 zu ermutigen. Dies kann sich auf das Weitermachen in  
298 sozialen Medien (z. B. weiteres Posten oder Liken von  
299 Beiträgen), auf die Stärkung von Vertrauen oder auf die  
300 Verringerung bzw. Beseitigung von Sorgen, Unsicherheit und  
301 ähnlichen Gefühlen beziehen.

302 (3) Ich finde das ist sehr Geil weil am Anfang ist das sehr  
303 gut also macht immer so weiter  
304 (NDY/179)

305 *Gratitude* – Ausdrücke aufrichtiger Dankbarkeit, die  
306 unaufgefordert bzw. spontan geäußert werden. Diese  
307 Dankesäußerungen sind deutlich abzugrenzen von  
308 routinierter Höflichkeit, wie reaktivem Danken (z. B. „Danke

---

übernommen. Das Sigle (NDY/xxx) zeigt, dass es sich um ein Kommentar zum Video xxx aus dem NottDeuYTSch-Korpus handelt. In Kommentaren, die über die eigentlichen Flauschdrücke hinausgehenden Text enthalten, werden die entsprechenden Flauschdrücke durch Unterstreichung hervorgehoben.

## Clausen & Scheffler: Flausch

309 für die Antwort“) oder formelhaften Ausdrücken wie „Danke  
310 im Voraus“.

311 (4) Danke , dass du das für uns machst 🥰❤️💖  
312 (NDY/033)

313 *Group membership* – Ausdrücke, mit denen eine bestimmte  
314 Gruppenzugehörigkeit signalisiert wird, wie zum Beispiel zur  
315 Fangemeinschaft eines/r bestimmten Contentcreator\*in.  
316 Dazu zählen insbesondere Hashtags und Emojis, die als  
317 Markenzeichen der jeweiligen Creator\*innen oder als  
318 Erkennungszeichen der jeweiligen Community fungieren,  
319 aber auch explizite, wörtliche Bekundungen der  
320 Zugehörigkeit.

321 (5) Einhörner 🦄🦄🦄🦄🦄🦄🦄🦄🦄🦄🦄🦄  
322 #BonnyTrash<sup>7</sup>  
323 (NDY/202)

324 *Positive feedback* – Äußerungen, die eine positive Haltung  
325 gegenüber dem ursprünglichen Beitrag oder einem  
326 Kommentar ausdrücken, jedoch nicht gegenüber einer  
327 spezifischen Person. Auch Emojis und Emoticons, die ohne  
328 begleitende sprachliche Äußerungen gepostet werden, fallen  
329 in diese Kategorie.

330 (6) Könnt ihr mal vielleicht wieder ein fakten video über  
331 euch machen wäre mega geil und die parodie ist mega  
332 hammer geil 🥰🥰🥰🥰🥰  
333 (NDY/186)

334 *Sympathy* – Äußerungen, die Mitgefühl und Verständnis  
335 ausdrücken. Diese können als Reaktionen auf negative oder  
336 betrübliche Mitteilungen anderer Personen erfolgen, aber  
337 auch initiativ Sympathie ausdrücken, wie zum Beispiel in  
338 Bezug auf eine bestimmte Lage oder Situation, die die  
339 Zielperson betrifft.

340 (7) + Elma D. Oh nein ! 🤔 Aufwärmen ist so so wichtig !!!  
341 Ich hoffe es geht dir langsam besser !

<sup>7</sup> Zum Einhorn als Gruppen-Kennzeichen der BonnyTrash-Fangemeinde, siehe auch Cotgrove (2025, Kapitel 6).

342 (NDY/033)

343 Die folgenden drei Annotationskategorien beziehen sich auf  
344 Sonderfälle, die zwar nicht unmittelbar Flausch ausdrücken,  
345 jedoch für die Analyse als relevant erachtet und daher mit  
346 annotiert wurden.

347 *Implicit* – Flausch wird nicht direkt geäußert, sondern kann  
348 aus einem bestimmten wörtlichen Ausdruck abgeleitet  
349 werden. Damit ein Ausdruck als implizite Form von Flausch  
350 annotiert werden kann, muss eindeutig erkennbar sein,  
351 welchem der zuvor definierten Flausch-Typen er zuzuordnen  
352 ist.

353 (8) Wieso macht ihr nicht bei dsds oder bein supertalent  
354 mit ? ihr seit sooooo gut  
355 (NDY/003)

356 *Agreement* – Ausdruck von Zustimmung zu der Meinung oder  
357 Aussage einer anderen Person. Dabei werden ausschließlich  
358 Kommentare berücksichtigt, die sich nachweislich auf  
359 Flauschäußerungen beziehen.

360 (9) Bin ganz deiner Meinung :)  
361 (NDY/179)

362 *Ambiguous* – diese Kategorie wurde in den Fällen  
363 verwendet, in denen nicht eindeutig bestimmt werden  
364 konnte, ob es sich um eine Flauschäußerung handelt, zum  
365 Beispiel aufgrund fehlenden Kontexts oder mehrdeutiger  
366 Ausdrucksweise.

367 (10) OMG  
368 (NDY/247)

## 369 4 Empirische Analyse

### 370 4.1 Datengrundlage

371 Als Datengrundlage für die vorliegende Untersuchung dient  
372 das NottDeuYTSch-Korpus (Cotgrove, 2023). Das Korpus  
373 umfasst über 33 Millionen Wörter aus rund 3 Millionen  
374 YouTube-Kommentaren, die im Zeitraum von 2008 bis 2018  
375 von jungen, deutschsprachigen Nutzer\*innen verfasst

## Clausen & Scheffler: Flausch

376 wurden. Für unsere Analyse haben wir Kommentare zu  
377 insgesamt 16 Videos von sieben verschiedenen Content-  
378 Creator\*innen ausgewählt (s. Tabelle 1). Die zu diesen Videos  
379 geposteten Kommentare wurden hinsichtlich des  
380 Vorkommens von Flausch analysiert und gemäß der in  
381 Abschnitt 3.3 erläuterten Flauschtypen annotiert.<sup>8</sup> Im  
382 folgenden Abschnitt wird die Verteilung der annotierten  
383 Flauschtypen in den Daten näher beschrieben und diskutiert.

# Discussion Paper

---

8 Der annotierte Datensatz steht für Forschungszwecke frei zur Verfügung und kann via OSF abgerufen werden: <https://osf.io/efh7c/>.

## Clausen & Scheffler: Flausch

Videoauthor*in	Video-Themen	Dateiname	Anzahl annotierter Kommentare
Anna Maria Damm	Lifestyle, Beauty, Fashion, Pflege-Routinen	NDY/285	203
DieLochis	Musik, Songparodien, Unterhaltung	NDY/003 NDY/004 NDY/179 NDY/186 NDY/208	32 463
Nilam	Beauty, Lifestyle, Alltags- und Kreativ-Vlogs	NDY/020 NDY/138 NDY/250 NDY/252 NDY/274	8 385
SophiaThiel	Fitness, Bodybuilding, Ernährung	NDY/033	2 387
xKarenina	Schmink-Tutorials, Buch- und Produktvorstellungen	NDY/246	241
Bonnytrash	Lifestyle, Vlogs, Challenges	NDY/202 NDY/203	925
Sami Slimani	Mode, Styling, Beauty, Pflege	NDY/247	1 682
<b>Gesamt</b>			46 286

384 **Tabelle 1:** Übersicht der Datengrundlage.

### 385 4.2 Verteilung der Flauschtypen

386 Insgesamt wurden 46 286 Kommentare hinsichtlich  
 387 Flauschäußerungen analysiert. In 14 579 dieser Kommentare  
 388 konnte mindestens ein Flausch Ausdruck identifiziert werden;  
 389 in Summe wurden 21 785 Flauschäußerungen annotiert.  
 390 Tabelle 2 zeigt die Häufigkeitsverteilung der annotierten  
 391 Flauschtypen.

## Clausen & Scheffler: Flausch

	Flauschtyp										
Creator*in	positive feedback	affection declaration	compliment	encouragement	group membership	gratitude	sympathy	agreement	implicit	ambiguous	Gesamt
Anna Maria	55 (29,9%)	4 (2,2%)	43 (23,4%)	13 (7,1%)	1 (0,5%)	5 (2,7%)	59 (32,1%)	1 (0,5%)	2 (1,1%)	1 (0,5%)	184 (100%)
DieLochis	6131 (54,7%)	2592 (23,1%)	1124 (10%)	534 (4,8%)	170 (1,5%)	34 (0,3%)	15 (0,1%)	225 (2%)	161 (1,4%)	218 (1,9%)	11204 (100%)
Nilam	3529 (53,2%)	567 (8,5%)	1919 (28,9%)	300 (4,5%)	11 (0,2%)	205 (3,1%)	14 (0,2%)	26 (0,4%)	36 (0,5%)	28 (0,4%)	6635 (100%)
Sophia	690 (52,7%)	107 (8,2%)	183 (14%)	82 (6,3%)	3 (0,2%)	204 (15,6%)	3 (0,2%)	9 (0,7%)	25 (1,9%)	4 (0,3%)	1310 (100%)
xKarenin	70 (59,8%)	10 (8,5%)	17 (14,5%)	1 (0,9%)	0 (0%)	1 (0,9%)	9 (7,7%)	0 (0%)	2 (1,7%)	7 (6%)	117 (100%)
Bonnytrash	561 (36,9%)	462 (30,4%)	58 (3,8%)	49 (3,2%)	363 (23,9%)	2 (0,1%)	0 (0%)	6 (0,4%)	15 (1%)	5 (0,3%)	1521 (100%)
Sami	364 (44,9%)	192 (23,7%)	160 (19,7%)	30 (3,7%)	10 (1,2%)	21 (2,6%)	1 (0,1%)	3 (0,4%)	14 (1,7%)	16 (2%)	811 (100%)
Gesam	11400 (52,3%)	3934 (18,1%)	3504 (16,1%)	1009 (4,6%)	558 (2,6%)	472 (2,2%)	101 (0,5%)	270 (1,2%)	255 (1,2%)	279 (1,3%)	21782 (100%)

392 **Tabelle 2:** Verteilung der Flauschtypen in den annotierten Daten.

393 Die quantitative Analyse verdeutlicht eine ungleiche  
394 Verteilung der verschiedenen Flauschtypen. *Positive*



## Clausen & Scheffler: Flausch

395 *feedback* ist mit 52,3% der am häufigsten vertretene  
396 Flauschtyp und macht damit etwas mehr als die Hälfte aller  
397 annotierten Flauschausdrücke aus. Das liegt unter anderem  
398 daran, dass diese Kategorie Äußerungen „allgemeiner Art“  
399 umfasst, beispielsweise solche, die positive Gefühle oder  
400 Einstellungen ausdrücken, ohne ein spezifisches  
401 Bezugsobjekt (Person oder Beitrag) zu benennen, wie zum  
402 Beispiel „voll geil“ oder „cool“. Ebenfalls in diese Kategorie  
403 fallen Emoticons und Emojis, die häufig eigenständig – also  
404 ohne begleitende sprachliche Elemente – gepostet werden.  
405 Diese Form nonverbaler Flauschrede lässt sich mit geringem  
406 Aufwand einsetzen und tritt daher besonders häufig auf. Sie  
407 stellt die visuelle Nachbildung typisierter Gesichtsausdrücke  
408 in grafischer Form dar und signalisiert eine „hohe emotionale  
409 Involviertheit“ (Beißwenger, 2015: 41). Emoticons und Emojis  
410 mit positiver Konnotation gelten als „gesichtsbestärkende  
411 Mittel“ (Beißwenger, 2019) und können daher als Flausch  
412 klassifiziert werden. In unseren Daten machen sie rund 2,4%  
413 der Äußerungen innerhalb der Kategorie *positive feedback*  
414 aus.

415 Die zweithäufigsten Flauschtypen sind solche, die sich auf  
416 Personen beziehen, beispielsweise durch positive  
417 Bewertungen persönlicher Eigenschaften und Leistungen  
418 (*compliment*; 16,1%) oder durch direkte Ausdrücke der  
419 Bewunderung und Zuneigung (*affection declaration*; 18,1%).  
420 Zusammengefasst machen diese personenbezogenen  
421 Formen des Flauschs etwa 34% aller annotierten Äußerungen  
422 aus und treten somit etwas seltener auf als weniger  
423 spezifizierte positive Rückmeldungen.

424 Die übrigen Flauschtypen machen jeweils weniger als fünf  
425 Prozent aller Flauschäußerungen aus: *encouragement* (4,6%),  
426 *group membership* (2,6%), *gratitude* (2,2%) und *sympathy*  
427 (0,5%). Die Kategorie *encouragement* dient primär dazu,  
428 YouTuber\*innen zum Durchhalten oder Weitermachen zu  
429 motivieren. Obwohl dieser Flauschtyp semantisch relativ  
430 allgemein gestaltet ist, tritt er seltener auf als die zuvor  
431 beschriebenen Kategorien. Dies könnte darauf  
432 zurückzuführen sein, dass er häufig an spezifische Situationen  
433 gebunden ist und daher weniger universell einsetzbar ist.  
434 Insbesondere findet er sich in Kommentaren zu Videos,  
435 deren Inhalte diese Art unterstützender Botschaften  
436 begünstigen, beispielsweise bei angehenden Sänger\*innen.

## Clausen & Scheffler: Flausch

437 Auch andere Kommentator\*innen können adressiert und zu  
438 positiver Interaktion, wie zum Beispiel dem Geben von Likes,  
439 ermutigt werden. Die anderen Flauschtypen sind inhaltlich  
440 enger gefasst. So signalisiert *group membership* die  
441 Zugehörigkeit zu bestimmten Communities, wie Fangruppen  
442 oder thematisch-orientierten Interessengemeinschaften. In  
443 unseren Korpusdaten verwenden Fans von Sänger\*innen  
444 DieLochis Hashtags wie #Lochinator und  
445 Selbstbezeichnungen wie „größter Fan“, um ihre  
446 Gruppenzugehörigkeit auszudrücken. Creator\*innen können  
447 spezifische Emojis gezielt einsetzen, um diese als  
448 Erkennungssymbole ihrer Community zu etablieren, wie zum  
449 Beispiel das 🌈-Emoji, das im Fall der Creatorin BonnyTrash  
450 als Gruppenidentitätssymbol fungiert (s. Beispiel 5). Auch die  
451 Kategorien *gratitude* und *sympathy* sind spezifischer  
452 ausgeprägt und treten nahezu ausschließlich in Kommentaren  
453 zu Videos auf, die entsprechende Emotionen hervorrufen. So  
454 finden sich Sympathiebekundungen in unseren Korpusdaten  
455 vor allem in Kommentaren zu einem Spendenaufruf für die  
456 Philippinen der YouTuberin Anna Maria Damm, während  
457 aufrichtiger Dank häufig in Kommentaren zu einem Fitness-  
458 Tutorial von Sophia Thiel geäußert wird. Die Kategorie  
459 *agreement* (1,2%) umfasst Flauschausdrücke, die als  
460 zustimmende Reaktionen auf flauschenthaltende  
461 Kommentare anderer Nutzer\*innen fungieren und somit  
462 Teile unterstützender Diskurse markieren. Da jedoch nicht  
463 alle Gesprächsverläufe vollständig rekonstruiert werden  
464 konnten (z. B. aufgrund gelöschter Accounts oder nicht mehr  
465 zuordenbarer Antwortbezüge), konnte diese Kategorie nur  
466 eingeschränkt erfasst werden.

### 467 **5 Klassifikationen positiver und negativer Sprache im** 468 **Vergleich**

469 Wie bereits erwähnt, stellt Flauschrede – ebenso wie andere  
470 Formen positiver Sprache – den Gegenpol zur Hassrede in  
471 der virtuellen Kommunikation dar. Während Hassrede darauf  
472 abzielt, Personen zu verletzen und deren Selbstbild zu  
473 destabilisieren, zielt positive Sprache darauf ab, das Selbstbild  
474 der Adressat\*innen aufzubauen, zu stärken und zu  
475 unterstützen.

476 Im Folgenden beleuchten wir zunächst bestehende  
477 Klassifikationen positiver und negativer Sprache und führen

## Clausen & Scheffler: Flausch

478 anschließend eine direkte Gegenüberstellung mit unserer  
479 Klassifikation von Flausch durch.

### 480 5.1 Klassifikationen positiver Sprache

481 Die bislang entwickelten Ansätze sowie die verfügbaren  
482 Datensätze zur Analyse positiver Sprache – insbesondere zur  
483 *hope speech* und *empowering language* (s. Abschnitt 3.2) –  
484 beschränken sich auf binäre Klassifikationen von Social-  
485 Media-Beiträgen. Dabei wird lediglich erfasst, ob bestimmte  
486 Merkmale der jeweiligen Form positiver Sprache in einem  
487 Beitrag vorhanden sind oder nicht, während feinere  
488 Differenzierungen dabei unberücksichtigt bleiben.

489 So wird in der Studie von Njoo et al. (2023) danach  
490 klassifiziert, ob Beiträge auf Reddit als *empowering* gelten  
491 oder nicht. Auch Ansätze zur *hope speech* folgen überwiegend  
492 diesem binären Schema und kategorisieren vollständige  
493 Beiträge auf verschiedenen Social-Media-Plattformen  
494 hinsichtlich des Vorliegens entsprechender Merkmale (s. u. a.  
495 Chakravarthi 2020; Malik et al. 2023; Nath 2025). Einige  
496 Arbeiten nehmen differenziertere Klassifikationen vor und  
497 unterscheiden zwischen verallgemeinerter, realistischer und  
498 unrealistischer Hoffnung (Balouchzahi et al. 2023) oder  
499 zwischen Varianten von Hoffnungsrede, die Glaube, Wunsch  
500 oder Kontrafaktizität ausdrücken (da Silva et al. 2025).  
501 Dennoch bleibt auch hier der gesamte Kommentar oder Post  
502 die zentrale Analyseeinheit.

503 Wie unsere Analyse der YouTube-Daten zeigt, sind  
504 Kommentare jedoch häufig sprachlich ambivalent und lassen  
505 sich nicht eindeutig als positiv oder negativ klassifizieren. So  
506 tritt Flausch oft gemeinsam mit abwertenden oder neutralen  
507 Formulierungen auf. Um dieser Komplexität gerecht zu  
508 werden, ist eine feingranulare Differenzierung notwendig,  
509 sowohl hinsichtlich verschiedener Flaushtypen als auch in  
510 Bezug auf die präzise sprachliche Abgrenzung relevanter  
511 Textsegmente innerhalb eines Beitrags. An genau diesem  
512 Punkt setzt der hier vorgestellte Ansatz an.

### 513 5.2 Klassifikationen von Hassrede

514 Hassrede ist ein komplexer und vielschichtiger Phänomen  
515 (Poletto et al. 2021: 478). In der Folge existieren weder eine  
516 einheitliche Terminologie für negative Onlinesprache noch

## Clausen & Scheffler: Flausch

eine konsistente Taxonomie von Hassrede-Ausdrücken (s. u. a. Schmidt/Wiegand 2017; Fortuna/Nunes 2018; Poletto et al. 2021; Jaki 2023). Neben dem Begriff „Hassrede“ finden sich in der Fachliteratur weitere Bezeichnungen wie *abusive language* (verletzende Sprache; Fortuna et al. 2023), *toxic language* oder *toxicity* (toxische Sprache; Pavlopoulos et al. 2020), *offensive language* bzw. *offensiveness* (beleidigende Sprache bzw. Anstößigkeit; Aggarwal et al. 2019; Zampieri et al. 2019) sowie *harmful speech* (schädigende Sprache; Scheffler/Solopova/Popa-Wyatt 2021; 2024).

Poletto et al. (2021) zeigen, dass diese Konzepte zwar eng miteinander verwoben sind, sich jedoch nicht eins zu eins aufeinander abbilden lassen. Ein vergleichbares Bild ergibt sich in Abschnitt 3.2 im Hinblick auf Flauschrede und verwandte Phänomene: Auch im Bereich positiver expressiver Sprache existieren unterschiedliche Konzepte, die bestimmte Teilaspekte abbilden, sich jedoch in ihren Einzelheiten unterscheiden. Wie im Bereich positiver Sprache sind auch die meisten Klassifikationen von Hassrede und verwandten Phänomenen binär angelegt und bestimmen, ob eine Aussage Hassrede enthält oder nicht. Einige Ansätze erweitern dieses Schema um zusätzliche Klassen, beispielsweise durch die Unterscheidung zwischen obszönen (*obscene*) und beleidigenden (*offensive*) Hassausdrücken (Mubarak et al. 2017) oder zwischen Hassrede und Ausdrücken, die beleidigend sind, aber nicht als Hassrede gelten (*offensive but not hate speech*) (Martins et al. 2018). Darüber hinaus existieren Ansätze, die graduelle Abstufungen von Hassrede berücksichtigen, wie zum Beispiel sehr toxisch, toxisch, leicht toxisch und nicht toxisch (Kolhatkar et al. 2019) oder direkt aggressiv, indirekt aggressiv und nicht aggressiv (Kumar et al. 2018). Zudem werden häufig adressatenbezogene Differenzierungen vorgenommen, unter anderem entlang sozialer, religiöser, sexistischer oder rassistischer Dimensionen (Waseem/Hovy 2016) sowie zwischen aufhetzender, kommunaler und religiöser Hassrede (Ishmam/Sharmin 2019).

Wenige Klassifikationen ermöglichen eine feingliedrigere Typisierung. Ein Beispiel hierfür ist der SwissText 2018 Hate Speech Hackathon Workshop, in dem Klassifikatoren entwickelt wurden, die eine Differenzierung zwischen toxisch, schwer toxisch, obszön, beleidigend, bedrohlich

## Clausen & Scheffler: Flausch

sowie Identitätshass ermöglichen. Eine ähnliche Herangehensweise findet sich im GermEval 2019 Shared Task zur Identifikation beleidigender Sprache, der zwischen Beleidigung, Profanität und Missbrauch differenziert (Struß et al. 2019). Schließlich existiert eine fünfteilige Taxonomie schädigender Sprache, die Unterkategorien umfasst (Scheffler/Solopova/Popa-Wyatt 2021). Diese Taxonomie ist nach ähnlichen funktionalen Aspekten aufgebaut wie die hier vorgeschlagene Kategorisierung von Flauschausdrücken und umfasst aufhetzende Sprache, abwertende Worte und Ausdrücke, beleidigende, missbräuchliche oder anstößige Verwendungen, ein- und ausgrenzende Sprache (In-/Out-Group) sowie Codes.

Unsere Klassifikation der Flausch-Ausdrücke (s. Abschnitt 3.3) umfasst sieben Typen, die unterschiedliche sprachliche Strategien unterstützender Kommunikation abbilden. Ergänzend beinhaltet sie zusätzliche Kategorien für impliziten Flausch (*implicit*), Zustimmung (*agreement*) sowie mehrdeutige Ausdrücke (*ambiguous*), die aufgrund fehlenden Kontexts nicht eindeutig als Flausch klassifiziert werden können.

Flausch und Hass lassen sich – in all ihren Ausprägungen – als zwei Seiten expressiver Sprache verstehen. Durch die Ausarbeitung differenzierter Annotationsrichtlinien für virtuelle Flausch- und Hassrede wird es möglich, empirische Untersuchungen zur Expressivität der Onlinesprache auf Basis von korpus- und computerlinguistischer Methoden durchzuführen (Scheffler 2025).

## 6 Zusammenfassung

In virtuellen Räumen finden vielfältige Formen sprachlicher und nicht-sprachlicher Interaktion statt. Als Gegenpol zum viel untersuchten Konzept der Hassrede definierten wir in diesem Beitrag den Begriff des virtuellen Flauschs. Darunter verstehen wir sprachliche Ausdrücke sowie nicht-sprachliche Elemente wie Emojis und Hashtags, die der Unterstützung und Stärkung des positiven Selbstbildes der Adressat\*innen dienen. Unsere Klassifikation unterscheidet sieben Typen von Flausch, darunter Komplimente, positives Feedback und Sympathiebekundungen, die sich durch spezifische linguistische Merkmale identifizieren lassen.



## Clausen & Scheffler: Flausch

Im Rahmen einer empirischen Untersuchung analysierten wir die Häufigkeit von Flauschausdrücken in Kommentaren zu deutschsprachigen YouTube-Videos aus dem NottDeuYTSch-Korpus. Etwa ein Drittel der Kommentare enthält Äußerungen, die unserer Definition von Flausch entsprechen. Diese Ergebnisse legen nahe, dass Flausch in bestimmten virtuellen Gemeinschaften deutlich präsenter ist als Hassrede. Aus unserer Sicht kann dies als ein möglicher Erklärungsansatz für die Attraktivität und Beliebtheit digitaler Kommunikation gelten.

Das Konzept der Flauschrede bietet eine vielversprechende Grundlage für zukünftige (computer-)linguistische Untersuchungen, die sich nicht nur auf negative, sondern explizit auch auf positive Interaktionen in virtuellen Lebenswelten fokussieren. Diese Studie verfolgte das Ziel, das Forschungsinteresse an der datenbasierten Analyse positiver expressiver Sprache zu fördern und neue methodische Perspektiven auf gesichtsunterstützende sprachliche Praktiken in virtuellen Räumen zu eröffnen.

## Literatur

- Aggarwal, Piush/Horsmann, Tobias/ Wojatzki, Michael/ Zesch, Torsten. (2019): LTL-UDE at SemEval-2019 Task 6: BERT and Two-Vote Classification for Categorizing Offensiveness. In *Proceedings of the 13th International Workshop on Semantic Evaluation*, 678–682, Minneapolis, Minnesota, USA. Association for Computational Linguistics. URL: <https://aclanthology.org/S19-2121/>
- Ashraf, Noman/Zubiaga, Arkaitz/Gelbukh, Alexander (2021): Abusive language detection in YouTube comments leveraging replies as conversational context. *PeerJ Computer Science*. DOI: [10.7717/peerj-cs.742](https://doi.org/10.7717/peerj-cs.742)
- Balouchzahi, Fazlourrahman/Sidorov, Grigori/Gelbukh, Alexander (2023): PolyHope: Two-level hope speech detection from tweets, Expert Systems with Applications, 225 (120078), 1-13. DOI: [10.1016/j.eswa.2023.120078](https://doi.org/10.1016/j.eswa.2023.120078).
- Beißwenger, Michael (2015): Sprache und Medien: Digitale Kommunikation. In *Studikurs Sprach- und Textverständnis*. Hypermediales E-Learning-Angebot des Ministeriums für Innovation, Wissenschaft und Forschung (MIWF) des Landes Nordrhein-Westfalen. URL:



## Clausen & Scheffler: Flausch

- 639 [http://www.michael-](http://www.michael-beisswenger.de/pub/beisswenger_digikomm_preview.pdf)  
640 [beisswenger.de/pub/beisswenger\\_digikomm\\_preview.pdf](http://www.michael-beisswenger.de/pub/beisswenger_digikomm_preview.pdf)  
641 Beißwenger, Michael/Pappert, Steffen (2019): *Face work* mit  
642 Emojis. Was linguistische Analysen zum Verständnis  
643 sprachlichen Handelns in digitalen Lernumgebungen  
644 beitragen können. In: Beißwenger, Michael/Knopp,  
645 Matthias (Hg.): *Soziale Medien in Schule und Hochschule:*  
646 *Linguistische, sprach- und mediendidaktische*  
647 *Perspektiven*. Frankfurt: Peter Lang (Forum Angewandte  
648 Linguistik 63), 101-143. URL:  
649 <https://www.peterlang.com/view/title/68195>  
650 Bendel, Oliver (2019): *350 Keywords Digitalisierung*. Springer  
651 Gabler, Wiesbaden. DOI: [10.1007/978-3-658-25823-8\\_3](https://doi.org/10.1007/978-3-658-25823-8_3)  
652 Brodnig, Ingrid (2016): *Hass im Netz: Was wir gegen Hetze,*  
653 *Mobbing und Lügen tun können*. Wien: Christian  
654 Brandstätter Verlag.  
655 Burnett, Heather (2017): Sociolinguistic interaction and  
656 identity construction: The view from game-theoretic  
657 pragmatics. *Journal of Sociolinguistics* 22 (1), 238–271.  
658 DOI: [10.1111/josl.12229](https://doi.org/10.1111/josl.12229)  
659 Burnett, Heather (2019): Signalling games, sociolinguistic  
660 variation and the construction of style. *Linguistics and*  
661 *Philosophy* 42 (5), 419–450. DOI: [10.1007/s10988-018-](https://doi.org/10.1007/s10988-018-9254-y)  
662 [9254-y](https://doi.org/10.1007/s10988-018-9254-y)  
663 Chakravarthi, Bharathi Raja (2020): HopeEDI: A Multilingual  
664 Hope Speech Detection Dataset for Equality, Diversity,  
665 and Inclusion. In: *Proceedings of the Third Workshop on*  
666 *Computational Modeling of People's Opinions, Personality,*  
667 *and Emotion's in Social Media*. Barcelona, Spain (Online).  
668 Association for Computational Linguistics, 41–53. URL:  
669 <https://aclanthology.org/2020.peoples-1.5/>  
670 Chakravarthi, Bharathi Raja/Muralidaran,  
671 Vigneshwaran/Priyadharshini, Ruba/Cn,  
672 Subalalitha/McCrae, John/García, Miguel Ángel/Jiménez-  
673 Zafra, Salud María/Valencia-García, Rafael/Kumaresan,  
674 Prasanna/Ponnusamy, Rahul/García-Baena, Daniel/García-  
675 Díaz, José (2022): Overview of the Shared Task on Hope  
676 Speech Detection for Equality, Diversity, and Inclusion. In  
677 *Proceedings of the Second Workshop on Language*  
678 *Technology for Equality, Diversity and Inclusion*. Dublin,  
679 Ireland. Association for Computational Linguistics, 378–  
680 388, DOI: [10.18653/v1/2022.ltedi-1.58](https://doi.org/10.18653/v1/2022.ltedi-1.58)

## Clausen & Scheffler: Flausch

- 681 Cotgrove, Louis (2023): THE NOTTDEUYTSCH CORPUS: A  
682 corpus of German-language YouTube comments, *Korpora*  
683 *Deutsch als Fremdsprache* 3 (2), 225–229. DOI:  
684 [10.48694/kordaf.3841](https://doi.org/10.48694/kordaf.3841)
- 685 Cotgrove, Louis (2025): *Abogeil! The language of German*  
686 *teens on YouTube*. Leibniz-Institut für Deutsche Sprache,  
687 Mannheim. DOI: [10.14618/amades-63](https://doi.org/10.14618/amades-63)
- 688 Da Silva, Tulio Ferreira Leite/Aduna, Gonzalo  
689 Freijedo/Benamara, Farah/Mari, Alda/Li, Zongmin/Yue,  
690 Li/Su, Jian (2025): CDB: A Unified Framework for Hope  
691 Speech Detection Through Counterfactual, Desire and  
692 Belief. In: *Findings of the Association for Computational*  
693 *Linguistics: NAACL 2025*, Albuquerque, New Mexico.  
694 Association for Computational Linguistics, 4448–4463.  
695 DOI: [10.18653/v1/2025.findings-naacl.252](https://doi.org/10.18653/v1/2025.findings-naacl.252)
- 696 García-Baena, Daniel/García-Cumbreras, Miguel  
697 Ángel/Jiménez-Zafra, Salud María/García-Díaz, José  
698 Antonio/Valencia-García, Rafael (2023): Hope speech  
699 detection in Spanish. *Lang Resources & Evaluation* 57,  
700 1487–1514. URL: [https://doi.org/10.1007/s10579-023-](https://doi.org/10.1007/s10579-023-09638-3)  
701 [09638-3](https://doi.org/10.1007/s10579-023-09638-3)
- 702 Goffman, Erving (1967): *Interaction Ritual: Essays on Face-*  
703 *to-Face Behavior*. Anchor Books, Doubleday & Company,  
704 Inc, Garden City, New York
- 705 Ishmam, Alvi Md/Sharmin, Sadia (2019): Hateful Speech  
706 Detection in Public Facebook Pages for the Bengali  
707 Language. *18th IEEE International Conference On*  
708 *Machine Learning And Applications (ICMLA)*, Boca Raton,  
709 FL, USA, 555-560. DOI: [10.1109/ICMLA.2019.00104](https://doi.org/10.1109/ICMLA.2019.00104)
- 710 Jaki, Sylvia (2023): Hate Speech in sozialen Medien: Ein  
711 Forschungsüberblick aus Sicht der Sprachwissenschaft. In:  
712 Jaki, Sylvia/Steiger, Stefan (Hg.), *Digitale Hate Speech:*  
713 *Interdisziplinäre Perspektiven auf Erkennung,*  
714 *Beschreibung und Regulation*. Berlin: J.B. Metzler.
- 715 Fortuna, Paula/Nunes, Sérgio (2018): A survey on automatic  
716 detection of hate speech in text. *ACM Computing Surveys*  
717 *(CSUR)*, 51 (4), 1–30. DOI: [10.1145/3232676](https://doi.org/10.1145/3232676)
- 718 Fortuna, Paula/Juan Soler-Company/Leo Wanner (2023):  
719 Dataset annotation in abusive language detection. In  
720 Strippel, Christian/

## Clausen & Scheffler: Flausch

- 721 Paasch-Colberg, Sünje/Emmer, Martin/Trebbe, Joachim  
722 (Hg.), *Challenges and perspectives of hate speech research*,  
723 443–464. Berlin. DOI: [10.48541/dcr.v12.26](https://doi.org/10.48541/dcr.v12.26)
- 724 Hosseinmardi, Homa/Mattson, Sabrina Arredondo/Ibn Rafiq,  
725 Rahat/ Han, Richard/Lv, Qin/Mishra, Shivakant (2015):  
726 Analyzing Labeled Cyberbullying Incidents on the  
727 Instagram Social Network. In: Liu, TY./ Scollon, C./Zhu, W.  
728 (Hg.) *Social Informatics. SocInfo 2015. Lecture Notes in*  
729 *Computer Science* 9471. Springer, Cham. URL:  
730 [https://link.springer.com/chapter/10.1007/978-3-319-](https://link.springer.com/chapter/10.1007/978-3-319-27433-1_4)  
731 [27433-1\\_4](https://link.springer.com/chapter/10.1007/978-3-319-27433-1_4)
- 732 Kersten, Saskia/Lotze, Netaya (2019): Creating a Self-Image:  
733 Face-Work and Identity Construction Online. *Journal für*  
734 *Medienlinguistik*, 2 (2), 123–156. DOI:  
735 [10.21248/jfml.2019.20](https://doi.org/10.21248/jfml.2019.20)
- 736 Klüwer, Tina (2011): “I like your shirt” – Dialogue Acts for  
737 Enabling  
738 Social Talk in Conversational Agents. *Lecture Notes in*  
739 *Computer Science* 6895, 14–27.
- 740 Kolhatkar, Varada/Wu, Hanhan/Cavasso, Luca/Shukla,  
741 Kavan/Taboada, Maite (2020): The SFU Opinion and  
742 Comments Corpus: A Corpus for the Analysis of Online  
743 News Comments. *Corpus Pragmatics* 4, 155–190. DOI:  
744 [10.1007/s41701-019-00065-w](https://doi.org/10.1007/s41701-019-00065-w)
- 745 Kumar, Ritesh/Ojha, Atul Kr./Malmasi, Shervin/Zampieri,  
746 Marcos (2018): Benchmarking Aggression Identification in  
747 Social Media. In: *Proceedings of the First Workshop on*  
748 *Trolling, Aggression and Cyberbullying (TRAC-2018)*,  
749 Santa Fe, New Mexico, USA. Association for  
750 Computational Linguistics, 1–11. URL:  
751 <https://aclanthology.org/W18-4401/>
- 752 Malik, Muhammad Shahid Iqbal/Nazarova, Anna/Jamjoom,  
753 Mona Mamdouh/Ignatov, Dmitry I. (2023): Multilingual  
754 hope speech detection: A Robust framework using transfer  
755 learning of fine-tuning RoBERTa model. In: *Info & Claims*,  
756 Band 35 (8). DOI: [10.1016/j.jksuci.2023.101736](https://doi.org/10.1016/j.jksuci.2023.101736)
- 757 Martins, Ricardo/Gomes, Marco/Almeida, Jose Joao/Novais,  
758 Paulo/ Henriques, Pedro (2018): Hate Speech Classification  
759 in Social Media Using Emotional Analysis, *7th Brazilian*  
760 *Conference on Intelligent Systems (BRACIS)*, Sao Paulo,  
761 Brazil, 61–66, DOI: [10.1109/BRACIS.2018.00019](https://doi.org/10.1109/BRACIS.2018.00019)

## Clausen & Scheffler: Flausch

- 762 Mubarak, Hamdy/Darwish, Kareem/Magdy, Walid (2017):  
763 Abusive Language Detection on Arabic Social Media. In:  
764 *Proceedings of the First Workshop on Abusive Language*  
765 *Online*, Vancouver, BC, Canada. Association for  
766 Computational Linguistics, 52–56. DOI: [10.18653/v1/W17-](https://doi.org/10.18653/v1/W17-3008)  
767 [3008](https://doi.org/10.18653/v1/W17-3008)
- 768 Nath, Tanusree/Singh, Vivek Kumar/Gupta, Vedika (2025):  
769 BongHope: An annotated corpus for Bengali hope speech  
770 detection. *International Journal of Information Technology*  
771 17, 2523–2531. DOI: [10.1007/s41870-025-02484-2](https://doi.org/10.1007/s41870-025-02484-2)
- 772 Njoo, Lucille/Park, Chan/Stappart, Octavia/Thielk,  
773 Marvin/Chu, Yi/Tsvetkov, Yulia (2023): TalkUp: Paving the  
774 Way for Understanding Empowering Language. In:  
775 *Findings of the Association for Computational Linguistics:*  
776 *EMNLP 2023*, Singapore. Association for Computational  
777 Linguistics, 9334–9354, DOI: [10.18653/v1/2023.findings-](https://doi.org/10.18653/v1/2023.findings-emnlp.625)  
778 [emnlp.625](https://doi.org/10.18653/v1/2023.findings-emnlp.625)
- 779 Palakodety, Shriphani/KhudaBukhsh, Ashiqur R./Carbonell,  
780 Jaime G. (2020): *Hope Speech Detection: A Computational*  
781 *Analysis of the Voice of Peace*. 24th European Conference  
782 on Artificial Intelligence – ECAI 2020. Santiago de  
783 Compostela, Spain. URL:  
784 [https://ecai2020.eu/papers/869\\_paper.pdf](https://ecai2020.eu/papers/869_paper.pdf)
- 785 John Pavlopoulos/Jeffrey Sorensen/Lucas Dixon/Nithum  
786 Thain/Ion Androutsopoulos. (2020): Toxicity Detection:  
787 Does Context Really Matter? In *Proceedings of the 58th*  
788 *Annual Meeting of the Association for Computational*  
789 *Linguistics*, 4296–4305, Online. Association for  
790 Computational Linguistics. DOI: [10.18653/v1/2020.acl-](https://doi.org/10.18653/v1/2020.acl-main.396)  
791 [main.396](https://doi.org/10.18653/v1/2020.acl-main.396)
- 792 Poletto, Fabio/Basile Valerio/Sanguinetti, Manuela/Bosco,  
793 Cristina/Patti, Viviana (2020): Resources and benchmark  
794 corpora for hate speech detection: a systematic review.  
795 *Lang Resources & Evaluation* 55, 477–523. DOI:  
796 [10.1007/s10579-020-09502-8](https://doi.org/10.1007/s10579-020-09502-8)
- 797 Rheingold, Howard (1994): *Virtuelle Gemeinschaft. Soziale*  
798 *Beziehungen im Zeitalter des Computers*. Bonn u.a.:  
799 ADDISON-WESLEY
- 800 Rieger, Stefan/Schäfer, Armin/Tuschling, Anna (2021):  
801 *Virtuelle Lebenswelten: Körper – Räume – Affekte*, Berlin,  
802 Boston: De Gruyter. DOI: [10.1515/9783110638127](https://doi.org/10.1515/9783110638127)



## Clausen & Scheffler: Flausch

- 803 Scheffler, Tatjana (2025): Expressivity and Computational  
804 Linguistics. In: Gutzmann, Daniel/Turgay, Katharina (Hg.):  
805 *The Oxford Handbook of Expressivity*. Oxford: Oxford  
806 University Press.
- 807 Scheffler, Tatjana/Solopova, Veronika/Popa-Wyatt, Michaela  
808 (2021): The Telegram Chronicles of Online Harm. *Journal*  
809 *of Open Humanities Data* 7 (8). DOI: [10.5334/johd.31](https://doi.org/10.5334/johd.31)
- 810 Scheffler, Tatjana/Solopova, Veronika/Popa-Wyatt, Michaela  
811 (2024): Verbreitungsmechanismen schädigender Sprache  
812 im Netz: Anatomie zweier Shitstorms. In: Gaderer,  
813 Rupert/Grömmke, Vanessa (Hg.), *Hass teilen: Tribunale*  
814 *und Affekte virtueller Streitwelten*. Bielefeld: transcript  
815 Verlag, 199–224. DOI: [10.1515/978383839471951-010](https://doi.org/10.1515/978383839471951-010)
- 816 Schmidt, Anna/Wiegand, Michael (2017): A survey on hate  
817 speech detection using natural language processing. In  
818 *Proceedings of the fifth international workshop on natural*  
819 *language processing for social media*. Valencia, Spain.  
820 Association for Computational Linguistics, 1–10. DOI:  
821 [10.18653/v1/W17-1101](https://doi.org/10.18653/v1/W17-1101)
- 822 Struß, Julia Maria/Siegel, Melanie/Ruppenhofer  
823 Josef/Wiegand, Michael/Klenner, Manfred (2019):  
824 Overview of GermEval Task 2, 2019 Shared Task on the  
825 Identification of Offensive Language. In: *Proceedings of the*  
826 *15th Conference on Natural Language Processing*  
827 *(KONVENS 2019)*, 354–365. URL:  
828 [https://corpora.linguistik.uni-](https://corpora.linguistik.uni-erlangen.de/data/konvens/proceedings/papers/germeval/GermEvalSharedTask2019Iggsa.pdf)  
829 [erlangen.de/data/konvens/proceedings/papers/germeval/G](https://corpora.linguistik.uni-erlangen.de/data/konvens/proceedings/papers/germeval/GermEvalSharedTask2019Iggsa.pdf)  
830 [ermEvalSharedTask2019Iggsa.pdf](https://corpora.linguistik.uni-erlangen.de/data/konvens/proceedings/papers/germeval/GermEvalSharedTask2019Iggsa.pdf)
- 831 Suler, John. 2004: The Online Disinhibition Effect.  
832 *CyberPsychology & Behavior*, 7 (3), 321–326. DOI:  
833 [10.1089/1094931041291295](https://doi.org/10.1089/1094931041291295)
- 834 Waseem, Zeerak/Hovy, Dirk. 2016: Hateful Symbols or  
835 Hateful People? Predictive Features for Hate Speech  
836 Detection on Twitter. In: *Proceedings of the NAACL*  
837 *Student Research Workshop*, 88–93, San Diego, California.  
838 Association for Computational Linguistics. DOI:  
839 [10.18653/v1/N16-2013](https://doi.org/10.18653/v1/N16-2013)
- 840 Wiegand, Michael/Siegel, Melanie/Ruppenhofer, Josef (2019):  
841 Overview of the GermEval 2018 Shared Task on the  
842 Identification of Offensive Language. In: *Proceedings of*  
843 *GermEval 2018, 14th Conference on Natural Language*

Clausen & Scheffler: Flausch

844      *Processing (KONVENS 2018)*, Vienna, Austria, 1–10. URL:  
845      <https://www.austriaca.at/8435-5>  
846      Zampieri, Marcos/Malmasi, Shervin/Nakov,  
847      Preslav/Rosenthal, Sara / Farra, Noura/Kumar, Ritesh.  
848      2019: Predicting the Type and Target of Offensive Posts in  
849      Social Media. In *Proceedings of the 2019 Conference of the*  
850      *North American Chapter of the Association for*  
851      *Computational Linguistics: Human Language*  
852      *Technologies*, Band 1 (Long and Short Papers), 1415–1420,  
853      Minneapolis, Minnesota. Association for Computational  
854      Linguistics. URL: <https://aclanthology.org/N19-1144/>

Discussion Paper