

TEACHING GERMAN WITH YOUTUBE COMMENTS: The benefits of authentic corpus-based material for younger learners

Louis Cotgrove
Leibniz-Institut für Deutsche Sprache Mannheim

Abstract

Das NottDeuYTSch-Korpus ist eine frei verfügbare Sammlung von YouTube-Kommentaren, die von Jugendlichen zwischen 2008 und 2018 unter ausgewählten deutschsprachigen Videos geschrieben wurden. Dieser Artikel untersucht am Beispiel des NottDeuYTSch-Korpus wie sich aus YouTube-Kommentaren, Lernmaterialien entwickelt werden können und wie Korpora aus internetbasierter Kommunikation den Lernprozess von fortgeschrittener Deutschlernenden bereichern können. Der Artikel erläutert die Auswirkungen authentischer Kommunikation in den YouTube-Kommentaren auf jugendliche Lernende und untersucht insbesondere, wie sich die psycholinguistischen Faktoren Motivation, Fremdsprachenangst und „Willingness to Communicate“ durch den Korpuseinsatz beeinflussen lassen. Außerdem wird auf die Vor- und Nachteile der Verwendung authentischer Korpusmaterialien in der Unterrichtsgestaltung eingegangen.

Keywords: Jugend; Zweitspracherwerb; digitale Kommunikation; authentische Ressourcen; YouTube

Abstract

The NottDeuYTSch corpus is a freely available collection of YouTube comments written under German-speaking videos by young people between 2008 and 2018. The article uses the NottDeuYTSch corpus to investigate how YouTube comments can be used to produce learning materials and how corpora of Digitally-Mediated Communication can benefit intermediate learners of German. The article details the effects of authentic communication within YouTube comments on teenage learners, examining how they can influence the psycholinguistic factors of motivation, foreign language anxiety, and willingness to communicate. The article also discusses the benefits and limitations of using authentic corpus material for the development of teaching material.

Keywords: youth; second language acquisition; computer-mediated communication; authentic materials; YouTube

1. Introduction

Over the past 20 years, corpora and corpus-based resources have become increasingly important to the field of Second and Foreign Language Acquisition (SLA),¹ in particular by providing educators and learners access to “‘authentic’ language” (Mitchell 2020: 254). Despite “corpora [...] increasingly contributing to state-of-the art research methodologies in the field [of SLA]” (Deshors / Gries 2023: 164), shortcomings within existing resources remain. These include a lack of metadata on the contexts and source of the language, and speaker, and often cannot be used for longitudinal analysis (ibid.: 173), which can limit the usefulness of the data when producing new resources. Furthermore, existing corpo-

¹ “SLA” is used in this article due to its widespread usage in existing research, although the term is not without issue, in part due to its assumption of monolingual environments and learners (for a discussion on monolingualism in the classroom, see García / Leiva 2014).

ra tend to represent a relatively narrow range of text types and genres, a further limitation on their use, according to Paquot (2022: 36), who called for “more genuine input corpora”.

This article discusses a resource that addresses these gaps within corpus-driven linguistic material in SLA, namely the NottDeuYTSch corpus (cf. Cotgrove 2023). The NottDeuYTSch corpus contains over 33 million words taken from YouTube comments written under German-language videos targeted at young people between 2008 and 2018 (an in-depth explanation of the construction of the corpus can be found in Cotgrove 2022: 59-96). The NottDeuYTSch corpus provides one of the first corpus linguistic opportunities to study digitally-mediated communication (DMC) written by and between young people, containing a wealth of metadata that enables both text genre-specific and longitudinal research. The corpus can offer pragmatic, sociolinguistic, and psycholinguistic benefits for the teaching of German as an additional language (*Deutsch als Fremdsprache/Deutsch als Zweitsprache*, shortened to DaF/DaZ), in particular to teenage learners, one of the largest groups of learners of German, both in the German-speaking area and abroad, and can inform the development of a wide variety of linguistic resources.

The article is structured in two sections following the introduction. Following the introduction, Section 2 discusses the psycholinguistic challenges facing teenage learners of second or additional languages and why linguistic sources of DMC are well-positioned to address the needs of this group of learners. In particular, the section explains the current gap in existing corpus-based resources and introduces the NottDeuYTSch corpus.

Section 3 examines three broad areas of language that a teenage learner would experience in DMC (and offline), namely ‘non-standard’ language use, interactive situations, and how to index their identity as a young person, providing examples from the corpus of comments, in what areas of language learning they can be used, and how the learner would also psycholinguistically benefit from learning material containing these linguistic features. The section also discusses the potential limitations of using YouTube comments for teaching DaF/DaZ, and how educators can best take these into account.

2. Developing SLA materials

2.1 Psycholinguistic factors for teenage learners

Teenagers face particular challenges when learning languages in a classroom settings, partly due to the complex physiological, psychosocial, and cognitive developments associated with puberty and adolescence (cf. Harklau 2007: 639-640). One of the effects this can have is a high “affective filter” (Krashen 1982), the concept that various attitudinal and environmental factors can influence the process of SLA. A learner with a low affective filter is able to more easily comprehend pedagogical input and therefore become more competent in the language, as compared to a learner with a high affective filter, which prevents the same input from effectively contributing to competence. While the affective filter model may have been developed over 40 years ago at time of writing, Behney / Marsden (2020: 38) argued that it is still relevant to the field today. This section will examine three of these factors, motivation, foreign language anxiety (FLA), and willingness to communicate (WTC), the role they play in teenage SLA, and how corpus-based DMC can be used to combat them, thus reducing the affective filter.

Motivation is the “desire to achieve the goal, positive attitudes, and effort” (MacIntyre 2002: 46). It has long been positively linked with achievement in SLA (cf. Krashen 1982: 30; Dunn / Iwaniec 2022: 988), and has even been called the “strongest predictor of achievement” (Ushioda / Dörnyei 2012: 403), not just among SLA (for an examination of motivation in mathematics, see Michaelides et al. 2019). Motivation can come from both internal factors to the learner and external influences, such as the classroom, and there is significant interplay between the two (cf. Dörnyei 2009).

Foreign language anxiety is a “subjective feeling of tension, apprehension, nervousness, and worry [in] the language learning situation” (Horwitz / Horwitz / Cope 1986: 125), which impedes the ability to process information and produce language in a classroom setting (cf. Sachs / Baralt / Gurzynski-Weiss 2023), and is particularly prevalent in teenagers and younger learners of languages (cf. MacIntyre et al. 2003; Russell 2020). This is in part due to the heightened fear of exposure to ridicule from peers or the avoidance of displays of vulnerability associated with puberty (cf. Taylor 2013: 2).² These anxieties have been reported even to lead to “students discontinuing their study of foreign languages” (Dewaele 2013: 177).

The willingness to communicate (WTC) by a learner refers to the intention to use a foreign language in a communicative context and has been subject to significant investigation since it was first proposed as a concept in the 1990s (see MacIntyre et al. 1998).³ Increasing WTC is central to SLA, not only is it relevant for the development of language fluency (cf. Myhre / Fiskum 2021: 213) and proficiency (cf. Reinders / Wattana 2014: 102), but Dörnyei (2001: 52) refers to WTC as “the most central objective” for communication in a foreign language.

While there are several ways of addressing the three factors individually, some are impractical for time-pressed educators, such as personalised psychological evaluations to reduce perfectionism and improve self-confidence and stress-resistance (see Dewaele 2013). An aspect of common ground that has been shown in previous research to combat all three factors, is the development of learner-centred materials, i.e., adapting materials to be relevant to the lived experiences of the learners, in this case teenagers, with a focus on using digital environments and digitally-native texts. This should be no surprise as young people spend a considerable amount of time engaging with DMC and digitally-mediated content; according to Beisch / Koch (2022: 461-464), over 99% of young people (14-29) in Germany used the internet daily and almost 2.5 hours were spent every day on engaging with social media and DMC.

Teenagers and young adult learners have been regularly shown to benefit from innovative learning materials and environments. Digital environments have been widely demonstrated to play an important role in increasing WTC and decreasing FLA amongst teenagers as they are often familiar with such communication situations, leading to an increase in confidence, a willingness to request help, a reduction in anxiety, and an increase in fluency (cf. Reinders / Wattana 2014: 115; Kruk 2019: 17-21; Jauregi-Ondarra / Canto / Melchor-Couto 2022). These environments include course management systems (cf. Erbaggio et al. 2012), the comments section for news articles (cf. Marchand / Akutsu 2015), video games (cf. Peterson 2016; Kruk 2019), and a mix of multiple digital platforms (cf. Lee / Lu 2021). However, simply placing learners in a digital context may not automatically promote their

² “SLA” is used in this article due to its widespread usage in existing research, although the term is not without issue, in part due to its assumption of monolingual environments and learners (for a discussion on monolingualism in the classroom, see García / Leiva 2014).

³ 77% of German 14-19 year-olds described themselves as active users of YouTube (cf. Statista 2023a), slightly behind Instagram with 79% (cf. Statista 2023b).

WTC in an additional language. It is important to combine the context with the access to learning material that contains cultural similarities to the learner (cf. Ballis 2014: 220), such as DMC.

The NottDeuYTSch corpus can prove useful as part of the development of SLA resources to combat all three factors: increasing motivation and WTC, and decreasing FLA, as a source of DMC familiar to the target learners. Integrating YouTube comments in the classroom can expand the existing materials in an educator's arsenal, to be used alongside YouTube videos, which have also been demonstrated to increase motivation (cf. Kabooha / Elyas 2018: 78; Dizon 2022: 24). By utilising language and (digital) communication situations familiar with learners, learners may more readily engage with the material and communicate in the target language.

However, it is relevant to stress that while the motivation of teenage learners increases and FLA decreases when DMC material is introduced into classroom materials (cf. Ritchie / Black 2012), it is DMC that was authentically produced, rather than texts contrived to resemble DMC. Section 2.2 discusses the importance of using authentic materials to construct and compile SLA resources, focusing on DMC and sourcing materials from corpora.

2.2 Suitable materials for SLA

The use of authentic language for learning materials is based on the premise that learners benefit from analysing real-world language data to discover patterns and regularities in the language and so become more aware of realities of communication in the target language, which can improve SLA in many stages of the language learning process (see Gilquin / Granger 2022). Previously, language teaching relied on educators contriving examples, such as conversations (for practice dialogues in 16th century, see Villoria-Prieto / Suso López 2018), or learning by rote in the 19th and early 20th centuries (cf. Cohen 2018). From the 1970s, the benefits of authentic materials for learners have been shown to improve cognitive engagement in the target language to solidify acquisition (cf. Craik / Lockhart 1972) and later work demonstrated further psychological and affective benefits, linking learner exposure to authentic materials with an increase in WTC, thus lowering the affective filter (cf. Gilmore 2007).

However, while contrived examples and texts do not prepare learners for the reality of language, most authentic language has been “produced in order to communicate rather than to teach”, therefore, it is recommended that “a version of an original which has been simplified to facilitate communication” (Tomlinson 2012: 161) be used for learners. For example, learners can be asked to search for patterns and regularities, to identify collocations, to analyse the use of grammar and vocabulary in context, and to use the language data to create their own texts (cf. Gilquin / Granger 2022: 434-436).

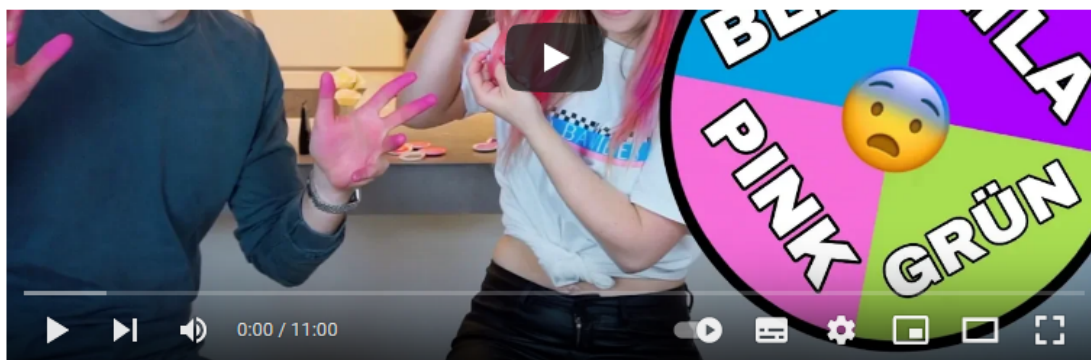
The use of DMC as a source of authentic language for SLA has significantly increased over the last 30 years, with studies employing a variety of DMC sources, including Internet Relay Chat (IRC) (cf. Smith 2005), online forums (cf. Meskill / Anthony 2005; Ritchie / Black 2012), email (cf. Vyatkina / Belz 2006), social media (cf. Chawinga 2017; Barrot 2022), online gaming (cf. Peterson 2016), watching YouTube videos (cf. Terantino 2011; Arndt / Woore 2018; Kabooha / Elyas 2018), and analysing YouTube and Facebook comments (cf. Chen 2020). While the majority of existing studies have investigated learners of English, research has also encompassed Russian (cf. Meskill / Anthony 2005), German (cf. Vyatkina / Belz 2006), and French (cf. Ritchie / Black 2012), and Terantino (2011) additionally offers a broad overview of resources for learning several languages. Studies have also demonstrated the benefits of using DMC in specific areas of SLA, for example, vocabulary learning (cf. Ka-

booha / Elyas 2018), writing skills (cf. Ritchie / Black 2012), modal particles in German (cf. Vyatkina / Belz 2006), and pragmatics (cf. Cunningham 2019).

Unfortunately, there are relatively few corpus-based resources of DMC for SLA. Existing SLA DMC corpora include the *Telekorp* corpus (cf. Vyatkina / Belz 2006) of email correspondence (roughly 1.5 million words) between learners of English and German, which serves as a useful bilingual corpus resource between the two languages. The *NAIST Lang-8* corpus (cf. Mizumoto et al. 2011) of language-learning social media posts (580,549 entries/comments) contains data in many languages, including English, Japanese, and Mandarin Chinese. The *deL1L2IM* corpus (cf. Höhn 2015) was compiled from instant messaging dialogues between German speakers and learners (52,000 tokens). The CMC learner and reference corpora (cf. Marchand / Akutsu 2015) were created from user comments underneath news websites (1.6m words). Finally, the *Uppsala WordReference* corpus (cf. Berdicevskis 2020) contains roughly 170m words taken from forum messages in English, Spanish, French, and Italian. However, none of the above corpora contain language explicitly produced by younger speakers that can be used to teach other younger speakers useful context-dependent language skills. This is representative of a wider problem, identified by Goschler / Stefanowitsch (2014), that there are few suitable corpora for learners that are representative of a particular group of speakers. Additionally, they identified three further problems with the corpus landscape that affects the ability of producing effective learning materials: many learner corpora are too small, they exhibit a lack of speaker variety, and it is often not clear what prompted the language production collected for the corpus data. Section 2.3 provides an explanation of how the NottDeuYTSch corpus can tackle these four limitations.

2.3 The NottDeuYTSch corpus

YouTube comments have been an, as yet, untapped source of authentic language from young people but studies of either DMC or youth language have rarely analysed the linguistic features used by young people in YouTube comments, despite the platform being the one of most-accessed means of communication in this demographic for several years (cf. Saferinternet.at 2018; Bahlo et al. 2019: 80) with 77% of German 14-19 year-olds describing themselves as active users of YouTube (cf. Statista 2023a), slightly behind Instagram with 79% (cf. Statista 2023b). The NottDeuYTSch corpus is one of the first large corpora of linguistic data containing online language from YouTube comments, i.e., what people have written underneath the videos, rather than the content of the video (see Figure 1).



GLÜCKSRAD entscheidet unsere HAARFARBE 🎲 | Bibi



BibisBeautyPalace ✓
5.84M subscribers

Subscribe

👍 75K



➦ Share

⌵ Save



2M views 4 years ago

JULIANS VIDEO 📺

Hier mich kostenlos abonnieren: <https://www.youtube.com/user/BibisBea...>

Show more

4,487 Comments ☰ Sort by



Add a comment...



@lavisworld3898 4 years ago

Könnt ihr mal bitte ein Bilou Haarshampoo machen? 🥰❤️

Ps: Wer das auch so sieht bitte liken damit sie das hoffentlich sehen 😊

👍 1.3K Reply

∨ 24 replies



@sugabape 4 years ago (edited)

Wie kann man seine Winterjacke verlieren?! 😂 Typisch Julian 😊

👍 2K Reply

∨ 97 replies

Figure 1

Screenshot of the comments section under a YouTube video from the popular German channel, *BibisBeautyPalace*

The NottDeuYTSch corpus can also address the four shortcomings mentioned in Section 2.2 in the corpus landscape for SLA learners in the following ways:

1. *A lack of suitable corpora for learners not representative of any particular group of speakers:* the corpus is a collection of authentically-produced DMC that has been specifically created to be representative of the language written under mainstream German YouTube channels by teenagers and young adults between

2008 and 2018. While the age of the commenters cannot be verified, a general age range of 12-21 can be inferred “based on the target demographic of the selected channels, reinforced by a post-hoc statistical analysis of explicit disclosures of age within the comments of the NottDeuYTSch corpus” (Cotgrove 2022: 62).

The corpus is primarily in German, although other languages such as English, Turkish and Serbian are also present, further representative of the multicultural digital linguistic landscape;

2. *Existing corpora are too small*: While there is no clear cut lower boundary for the size of a corpus, smaller corpora will not contain as many language features that appear less frequently, which may be useful to language learners. For a corpus to reliably contain enough syntactic features for analysis, it must contain at least one million tokens (see Baker 2010: 95-96). Containing over 33 million tokens, the NottDeuYTSch corpus is large enough to be used to analyse a significant range of linguistic features, and fulfills the call from Granger (2009, p. 28) for the implementation of longitudinal corpora in SLA, as it was created through stratified sampling based on the time stamp of the comment and genre of the video (see Cotgrove 2023);
3. *Lack of speaker variety*: The corpus contains over 1 million different speakers and discussion on a wide variety of topics, from beauty to gaming to current events. This allows text type analyses and comparisons, as well as broader generalisations;
4. *Unclear what prompted the language production collected for the corpus data*: an unclear or lack of reasoning for the collation of the corpus severely curtails pragmatic understanding of the language, reducing the usefulness of the authentic data, as the context of the language production is important for learners for comprehension and production (see also the Radiotelephony Plain English Corpus (RTPEC), specialised for learners of Aviation English Prado / Tosquil Lucks 2019; Friginal / Roberts 2022). The language in NottDeuYTSch corpus occurs either in comments directly responding to a video, or in replies to other comments, and contains metadata enabling conversational analysis.

A statistical breakdown of the corpus can be found in Table 1 (adapted from Cotgrove 2022: 343).⁴

⁴ For more on the development and modelling of WTC, see MacIntyre (2007) and Dörnyei (2003).

Statistic	Value
Number of Tokens (including emoji and emoticons)	33,760,494
Number of Tokens (only lexemes)	32,549,462
Number of Types	567,086
Type-Token Ratio (TTR)	0.017
Number of Comments	3,149,457
Number of Videos	296
YouTube Channels Represented	63
Mean Tokens per Comment	10.720
Median Tokens per Comment	5
Mean Comments per Video	1,914

Table 1
Statistical overview of the NottDeuYTSch corpus, adapted from Cotgrove (2022: 343)

YouTube itself has long been a platform containing SLA materials, in particular, content creators uploading videos that offer to teach languages, specifically in an extramural context (cf. Chowchong 2022). While these can be helpful to learners, the videos come with a host of caveats, such as the lack of quality control and reliable feedback method (cf. Dizon 2022). Within the YouTube platform, users are reliant on recommendations from the algorithm or reading judgements from other viewers in the comments to evaluate the potential quality of a video for SLA, especially since there is no easy way to quantitatively gauge quality following the removal of a public ‘dislike’ counter under a video. YouTube language content creators often try to use the platform to advertise paid courses, also without recognised qualification (cf. Bruzos 2021).

Within the comments section of SLA videos, there are often a significant amount of comments engaging in metalinguistic discourse, i.e., opining on ‘correct’ pronunciation or spelling, among other topics. While these comments have the potential to guide users further in their language learning, as well as providing a feedback mechanism if a commenter replies to a question, similar to an online forum, again quality control proves tricky for a learner; the number of comments containing metalinguistic discussion is often too large for the effective synthesis of knowledge when faced with so much new material (cf. Benson 2015). The applications for a corpus consisting of comments written under language videos would be limited for DaF/DaZ learners as there would be a lack of authentic target language, compounded by the heightened awareness of the commenters to write what they perceive as ‘correct’ German. While the NottDeuYTSch does contain metalinguistic discourse, it is produced in an environment where the viewers are not primed to think about adherence to lexicogrammatical rules.

Dizon (2022: 22) reported of concerns by learners that “[watching] YouTube [videos] was not an effective means to improve English ability when compared to more formal learning environments”,

despite finding YouTube “an interesting and flexible language learning method”. I would argue that teenage learners stand to significantly benefit from the combination of YouTube content, in this case comments, with the structure and feedback offered by teacher-led learning. Section 3 discusses how educators can create learning resources containing YouTube comments, providing several worked examples.

3. Using YouTube comments as a learning resource

The following section uses close analyses of several comments taken from the NottDeuYTSch corpus to detail several potential areas for activities and tasks for teenage learners, as well as potential limitations of the corpus. The areas for analysis have been developed following the guidelines of recent literature that has investigated the integration of corpus-based resources into learning resources, providing educators the option to e.g.:

1. Select specific corpora that cover particular text or genre types, e.g., digital communication, which can “focus the curriculum on domains of language which will be most useful to students” (Mitchell 2020: 254);
2. Draw from the ‘frequency and idiomaticity of language patterns in specific language varieties’ to produce resources (cf. Paquot 2022: 28);
3. Use corpus linguistic analysis to identify “formulaic sequences that are frequent and semantically transparent” as these “can serve as points of departure at intermediate stages of language development, especially when they are useful in a particular discourse situation or task” (Wulff 2020: 180).

Using corpora in this way would aid the comprehension of the text types present in DMC and enable the development of a wider variety of language tasks, that would also assist in increasing motivation and WTC, and reducing FLA for the teenage learners. The materials can be designed to target specific language features, such as collocations, idioms, and discourse markers. By focusing on these features, learners can develop a deeper understanding of the language and its use, and can improve their ability to use the language in a more naturalistic and effective way. For example, one situation where such focused lessons would be of particular benefit would be providing resources and exercises containing authentic online exchanges between young people ahead of an exchange programme for intermediate learners or for refugee learners to better integrate with their new peers.

The following sections address three different aspects of youth DMC in turn, using examples from the corpus as the basis for the potential development of learning materials by educators, namely ‘non-standard’ language use, interaction, and indexing youth identities.

3.1 ‘Non-standard’ language use

In the NottDeuYTSch corpus there are many comments containing examples where spelling, word order, punctuation, and other grammatical aspects would be considered as ‘non-standard’. Busch (2021) convincingly argued that the “binary paradigm of standard and non-standard” is outdated, but often within an institutionalised setting, such as a classroom, any deviation from a narrowly prescribed

‘standard’ language is classified as an error (cf. Milroy / Milroy 1999: 83), with the educator or learner unable or unwilling to discuss the nuances of linguistic variation. Therefore, a more pragmatic approach is necessary, where educators can use YouTube comments to demonstrate how young people communicate online, and how learners can identify and comprehend DMC-specific linguistic features, and how learners might expect the comment to be written in standard German. While non-standard language in DMC is not an exclusively youth-specific phenomenon, it occurs significantly more frequently and more varied in youth language than in other communicative situations, such as informal adult communication (for corpus-based comparisons of non-standard lexis and syntax, see Cotgrove 2022). Generally, non-standard language usage can be utilised by educators for a number of classroom activities including:

1. “Error analysis” (Corder 1975; James 2013), i.e., learners would identify linguistic features deemed non-standard;
2. Metalinguistic discussions on the texts (see Myhill / Newman 2016). These include:
 - a. Reflecting on what features of writing are appropriate for different settings (i.e., register);
 - b. What particular features are common or expected in these texts (i.e., style)?
 - c. What can be considered an error and why are they considered as such (i.e., metalinguistic discourse)?
3. Practice writing in these styles after reading the comments as part of a themed lesson, or as part of a larger exercise re-writing a set text in many different styles and registers.

The following examples taken from the corpus exhibit a wide range of non-standard features, which are discussed underneath:

- (1) ir seit einfach nur peinlich,unlustig und dum
(you are just embarrassing, unfunny and dumb)
- (2) die Frisur steht dir echt mega!
(The hair cut really mega suits you!)
- (3) Ich hoffe sehr das dass nur erfunden ist weil sonst wäre das echt krank...
(I seriously hope that that is only made up because otherwise that would be really sick)
- (4) Weißt wenn es gegongt hat und wir in die Pause wollen und dann der Lehrer auf einmal sagt der Lehrer beendet den Unterricht wofür is der Gong Dan bitte da :(
(Y’know when the gong went and we wanted to go to break and then the teacher suddenly says the teacher ends the lesson so why is the gong even there then :())

Example (1) contains orthographical variation in *ir* (*ihr*, which should also be capitalised at the beginning of the sentence), *seit* (*seid*), and *dum* (*dumm*), as well as a missing space following the comma between *peinlich* and *unlustig*. The comment would work well in discussing spellings and avoiding common mistakes based on phonology, such as *ihr seit*. The comment can also be used within material discussing interaction (see Section 3.2) and impoliteness strategies (see Lorenzo-Dus / Blitvich / Bou-Franch 2011).

Example (2) contains the phrase *echt mega*, an example of an increasingly common way for young people to express joy in informal situations, similar to *geil* in the last decades of the 20th century. The adverb *echt* serves as an opportunity to discuss intensification, such as alternatives to *sehr* when providing opinions. Furthermore, *mega* could serve as a springboard for discussion on language change, having originally been used as an intensifier itself, e.g., *megageil*, before first becoming an

interjection (*mega!*), and then an adjective in its own right, productive in adjectival phrases, and able to be intensified itself.

Example (3) also contains lexis specific to young people in *krank*, although in this comment used to express negative emotions. However, *krank*, similar to the (British) English *sick*, can also be used to express positivity (cf. Palacios Martínez 2018: 381), which can lead to discussions on meaning change. The comment also contains an example of a subordinate clause containing verb-second word order (*weil sonst wäre das echt krank*) instead of verb-last word order (*weil sonst das echt krank wäre*). The use of verb-second word order following subordinating conjunctions has dramatically increased over the past decades within youth language, roughly doubling from 2009 to 2018 (cf. Cotgrove 2022: 180). The phenomenon is not just restricted to *weil*; clauses starting with *obwohl*, *dass*, and *wenn* have all been found in the NottDeuYTSch corpus alone. This feature is useful for lessons on word order, which is often a sticking point for learners of German, especially for speakers of subject-verb-object languages, such as English. Other deviations from the standard include common ‘mistakes’, such as the confusion of *das* and *dass*, where *dass* is the subordinating conjunction and *das* the definite article, and the lack of commas separating the main clauses from the subordinating clauses.

The lack of commas separating clauses is also present in Example (4). The comment contains a good example in *wenn es gegongt hat* for discussing the differences between the uses of the conjunctions *wenn* (for conditional sentences), *als* (for temporal sentences), and *wann* (for questions) (even if the reality is not so clear cut, e.g., *wenn* in temporal case, see Breindl / Volodina / Waßner 2014: C1). The comment could be interpreted as particularly conceptually oral, i.e., it represents spoken language in written form, for example the pronoun-dropping (*weiß*, instead of *weiß du*). While a discussion on the applications of the *Nähe/Distanz* model (see Koch / Österreicher 2007) might exceed the capacity of a final-lesson-on-a-Friday-class of 15-16 year-olds, it might aid learners in exercises where they should write dialogue (also refer to Section 3.2).

3.2 Interaction

Authentic interaction between conversants is an often overlooked aspect of SLA, especially as there is a strong correlation between the quality of the input and the quality of a learner’s language output, with exposure to authentic language leading to effective production of the target language (see Li 2017: 58). Example (5) is a thread containing a short multi-party conversation between six participants.⁵ The video under which the comment thread was written, produced by comedy duo *DieAussenseiter*, has a deliberately low-budget aesthetic and a provocative style of humour that invites sarcastic or otherwise sarcastic comments.

- (5) **A:** Dieses Video ist schlecht! Muhahaha
B: Cool du spielst da mit
C: Du bist cool
D: +A Und du bist dumm! hahahahhhaahhahahahahaha!!!!
E: +B nein das macht er nicht

⁵ “SLA” is used in this article due to its widespread usage in existing research, although the term is not without issue, in part due to its assumption of monolingual environments and learners (for a discussion on monolingualism in the classroom, see García / Leiva 2014).

- E: +A und du bist ein opfa hahahahahahahahaha
 F: +A nur weil du selber schlecht bist musst du nicht andere belästigen 🖐️
- (A: This video is bad! Muhahaha
 B: Cool you are in the video too
 C: You are cool
 D: +A And you are dumb! hahahahahhhaahhahahahahaha!!!!
 E: +B No he's not
 E: +A and you are a loser hahahahahahahahaha
 F: +A just because you are bad doesn't mean you have to annoy others 🖐️)

This exchange can be used as a useful starting point for developing resources on a range of topics, such as orthography, interaction, and communication issues within DMC. Commenter A (all names have been anonymised) jokingly insults the video using extreme repetition of a representation of laughter (*Muhahaha*), to which Commenters D and E jokingly respond, calling A an *opfa* (*Opfer*, victim/loser) - a reference to a part of the video, and mirror the representations of laughter. However, Commenter F does not seem to understand the context and writes a comment that could be interpreted as a rejection of A (in particular, the 'middle finger' emoji) (for more on the role of emoji in YouTube comments, see Cotgrove 2022: 242). In the exchange, we see a demonstration of important pragmatic aspects of conversation between young people, such as the metacommunicative functions of laughter and repetition (see *ibid.*: 230). Here the exaggerated length of the representations of laughter (e.g., *hahahahahahaha-haha*) underlines the irony of the statements made by commenters A, D, and E, i.e., they are trying to ensure that other readers of their comments understand that they do not really mean the video is bad or another commenter is stupid (for more on irony in DMC, see Thompson et al. 2016). Learners can hope to also understand or identify with the role of insults and impoliteness, e.g., ritualised insults, that are particularly relevant for teenagers (see Kehily / Nayak 1997; for more on impoliteness, see Blitvich / Lorenzo-Dus / Bou-Franch 2013), or how to avoid misunderstandings when communicating online (see Bou-Franch / Lorenzo-Dus / Blitvich 2012), which can reduce FLA and increase WTC.

However, the thread, such as Example (5), does not necessarily need to be used verbatim for the development of teaching resources. Educators can adapt such materials to the needs of the class, as long as the basis is authentic material, as mentioned in Section 2.2. Here educators could, for example, ask learners to continue the conversation, either written or orally, in the same style, to practice turn-taking, conversation repair, and other aspects of interaction.

3.3 Indexing identities

The linguistic construction of youth identities has been a well-researched topic since the 1990s, often looking at constructions of youth ethnic identities (cf. e.g., Cutler 1999; Androutsopoulos 2005; Keim 2007; Kerswill 2013), but has tended to focus on lexis and phraseology, although more recent work has investigated grapheme-based variation (cf. Androutsopoulos 2018; Busch 2021), including emoji (cf. Cotgrove 2022: 251-252). Therefore, it is important to develop resources to increase WTC and motivation, and reduce FLA in teenage learners by focusing on how teenagers can use language to sound like an authentic teenager in the target language in DMC (and other youth-specific situations), in this case

corporation of YouTube in learning materials is not beneficial for spelling, although it was shown to improve vocabulary.

There is no guarantee that lessons based on YouTube would enthruse a group of teenagers; it might be seen as an attempt by the educator to ‘look cool’, which reduces, rather than increases WTC. The corpus might also not enthruse a wide range of educators, due to the linguistic variation in the comments, as both DMC-specific and youth-specific aspects may be considered to contravene the prescriptivist standard language ideologies present in many classrooms (cf. Milroy / Milroy 1999: 83), and as such is believed that it should not be taught to learners.

Furthermore, the NottDeuYTSch corpus was not created specifically be a learner corpus (for an overview of learner corpora, see Fernández / Davis 2021), although the texts were written by young people, learning German mostly as a first language (for a summary of the reasoning behind this, despite the lack of ethnographic data, see Cotgrove 2022: 62-64). This might require some extra annotation or markup from researchers or educators but there is the opportunity to use the NottDeuYTSch corpus as the basis for the creation of a learner corpus, e.g., for the analysis of parallels between learners of German as a first language and an additional language, or between DMC styles, such as with other existing corpora of online data, e.g., the *MoCoDa2* corpus of *WhatsApp* messages (cf. Beißwenger et al. 2020), the *DiDi* corpus of *Facebook* texts (cf. Glaznieks / Frey 2020), and the *Internetbasierte Kommunikation* (Internet-based communication, IBK) corpus, which contains blogs, emails, and internet relay chats (IRC) (cf. Lungen / Kupietz 2020).

The biggest restriction to the use of the corpus in learning material is linked to the rapidly changing nature of youth language and DMC. The earliest data in the corpus come from 2008, which means that it contains several linguistic features that are already ‘out of date’ for young people, e.g., the preference to use emoticons rather than emoji,⁶ the prevalence of *geil*, or pop culture references. That being said, the rate of change for syntactic elements is significantly slower than for lexis, meaning that the corpus will still be useful, for example in 10 years’ time for studies of word order, if not as a way of learning the latest slang. Furthermore, the production of such corpora take a considerable amount of time, from the research of the source material, to balancing, data extraction, and cleaning, meaning that no similar corpus can truly be current. The NottDeuYTSch corpus, on the other hand, is ready to be deployed.

4. Conclusion

The article has demonstrated the importance of using appropriate source material, in this case the NottDeuYTSch corpus of YouTube comments, to improve SLA for teenage learners, due to the high levels of exposure and familiarity teenagers have to YouTube and DMC environments. In particular, the corpus can be used to develop learning materials that use the authentic language produced by young people in YouTube comments to target three key affective factors, namely motivation, FLA, and WTC.

The NottDeuYTSch corpus can also be used to develop a pragmatic and metacommunicative understanding of DMC texts and interactions in teenage learners, to improve comprehension of ortho-

⁶ Emoticons are defined here as typographic approximations of faces and body parts rotated 90 degrees, primarily formed by combining punctuation and alphanumeric characters. Emoji are small graphical representations of faces, people, things, ideas, and concepts. designed and defined by the Unicode consortium (cf. Cotgrove 2022: 221-222).

graphical, lexical, and syntactic variation in youth language and DMC, and to acquire the linguistic resources that will allow them to express themselves in an age and context-appropriate manner, thus empowering the learner.

English still remains overwhelmingly dominant within research on DMC-based SLA (cf. Dizon 2022: 24) and the NottDeuYTSch corpus offers an opportunity for both educators and researchers to develop German-language-specific resources, such as learner corpus-oriented annotations, to advance corpus-based DaF/DaZ. It is hoped that the corpus can therefore help learners in their comprehension and production of German to fulfill the “ultimate goal of language learning”, according to MacIntyre et al. (1998: 559), namely, “authentic communication between persons of different languages and cultural backgrounds”.

Bibliography

Androutsopoulos, Jannis (2005): ... *und jetzt gehe ich chillen*: Jugend- und Szenesprachen als lexikalische Erneuerungsquellen des Standards. In: Eichinger, Ludwig / Kallmeyer, Werner (Eds.): *Standardvariation: Wie viel Variation verträgt die deutsche Sprache?* Berlin / Boston: de Gruyter, 171-206.

Androutsopoulos, Jannis (2018): Digitale Interpunktion: Stilistische Ressourcen und soziolinguistischer Wandel in der informellen digitalen Schriftlichkeit von Jugendlichen. In: Ziegler, Arne (Ed.): *Jugendsprachen: Aktuelle Perspektiven Internationaler Forschung*. Berlin / Boston: de Gruyter, 721-748.

Arndt, Henriette / Woore, Robert (2018): Vocabulary learning from watching Youtube videos and reading blog posts. In: *Language Learning and Technology* 22: 3. <https://ora.ox.ac.uk/objects/uuid:e5159356-78e6-4560-bd94-8d7242b256f8> (21.07.2023).

Bahlo, Nils / Becker, Tabea / Kalkavan-Aydin, Zeynep / Lotze, Netaya / Marx, Konstanze / Schwarz, Christian / Şimşek, Yazgül (2019): *Jugendsprache: Eine Einführung*. Berlin: J.B. Metzler.

Baker, Paul (2010): Corpus Methods in Linguistics. In: Litosseliti, Lia (Ed.): *Research Methods in Linguistics*. London: Continuum, 93-113.

Ballis, Anja (2014): Puschkin oder Podolski? – Schreiben in der Zweitsprache. In: Ahrenholz, Bernt / Grommes, Patrick (Eds.): *Zweitspracherwerb im Jugendalter*. Berlin / Boston: de Gruyter, 211-230.

Barrot, Jessie S. (2022): Social media as a language learning environment: a systematic review of the literature (2008-2019). In: *Computer Assisted Language Learning* 35: 9, 2534-2562.

Behney, Jennifer / Marsden, Emma (2020): Introduction to SLA. In: Paquot, Magali / Tracy-Ventura, Nicole (Eds.): *The Routledge handbook of second language acquisition and corpora*. London: Routledge, 37-49.

Beisch, Von Natalie / Koch, Wolfgang (2022): ARD/ZDF-Onlinestudie: Vier von fünf Personen in Deutschland nutzen täglich das Internet. In: *Media Perspektiven* 10, 460-469.

Beißwenger, Michael / Fladrich, Marcel / Imo, Wolfgang / Ziegler, Evelyn (2020): Die Mobile Communication Database 2 (MoCoDa 2). In: Marx, Konstanze / Lobin, Henning / Schmidt, Axel (Eds.): *Deutsch in Sozialen Medien: Interaktiv – multimodal – vielfältig*. Berlin / Boston: de Gruyter, 349-352.

Benson, Phil (2015): Commenting to Learn: Evidence of Language and Intercultural Learning in Comments on YouTube Videos. In: *Language Learning & Technology* 19: 3, 88-105.

- Berdicevskis, Aleksandrs (2020): Foreigner-directed speech is simpler than native-directed: Evidence from social media. In: Bamman, David / Hovy, Dirk / Jurgens, David / O'Connor, Brendon / Volkova, Svitlana (Eds.): *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science*. Association for Computational Linguistics, 163-172. <https://doi.org/10.18653/v1/2020.nlpcss-1.18>.
- Blitvich, Pilar Garcés-Conejos / Lorenzo-Dus, Nuria / Bou-Franch, Patricia (2013): Relational work in anonymous, asynchronous communication: A study of (dis)affiliation in YouTube. In: Kecskes, Istvan / Romero-Trillo, Jesús (Eds.): *Research Trends in Intercultural Pragmatics*. Berlin / Boston: de Gruyter, 343-366.
- Bou-Franch, Patricia / Lorenzo-Dus, Nuria / Blitvich, Pilar Garcés-Conejos (2012): Social interaction in YouTube text-based polylogues: A study of coherence. In: *Journal of Computer-Mediated Communication* 17: 4, 501-521.
- Breindl, Eva / Volodina, Anna / Waßner, Ulrich Hermann (2014): *Handbuch der deutschen Konnektoren 2, Semantik der deutschen Satzverknüpfers*. Berlin / Boston: de Gruyter. <https://doi.org/10.1515/9783110341447>.
- Bruzos, Alberto (2021): 'Language hackers': YouTube polyglots as representative figures of language learning in late capitalism. In: *International Journal of Bilingual Education and Bilingualism* 26: 10, 1-18.
- Busch, Florian (2021): Enregistered spellings in interaction: Social indexicality in digital written communication. In: *Zeitschrift für Sprachwissenschaft* 40: 3, 297-323.
- Chawinga, Winner Dominic (2017): Taking social media to a university classroom: teaching and learning using Twitter and blogs. In: *International Journal of Educational Technology in Higher Education* 14: 3.
- Chen, Cheryl Wei-Yu (2020): Analyzing online comments: a language-awareness approach to cultivating digital literacies. In: *Computer Assisted Language Learning* 33: 4, 435-454.
- Chowchong, Akra (2022): *Sprachvermittlung in den Sozialen Medien: eine soziolinguistische Untersuchung von DaF-Sprachlernvideos auf Videokanälen*. Berlin: Erich Schmidt Verlag.
- Cohen, Michèle (2018): From 'Glittering Gibberish' to 'Mere Jabbering'. In: McLelland, Nicola / Smith, Richard (Eds.): *The History of Language Learning and Teaching*. Cambridge: Legenda, 1-20.
- Corder, Stephen P. (1975): Error Analysis, Interlanguage and Second Language Acquisition. In: *Language Teaching* 8: 4, 201-218.
- Cotgrove, Louis Alexander (2018): Das Nottinghamer Korpus Deutscher YouTube-Sprache (the NottDeuYTSch corpus). LINDAT/CLARIAH-CZ. <http://hdl.handle.net/11372/LRT-4806>.
- Cotgrove, Louis Alexander (2022): #GlockeAktiv: A corpus linguistic investigation of German online youth language. Nottingham: University of Nottingham. <https://eprints.nottingham.ac.uk/id/eprint/69043> (21.07.2023).
- Cotgrove, Louis Alexander (2023): New opportunities for researching digital youth language: The NottDeuYTSch corpus. In: Kupietz, Marc / Schmidt, Thomas (Eds.): *Neue Entwicklungen in der Korpuslandschaft der Germanistik*. Tübingen: Narr, 101-114.
- Craik, Fergus I.M. / Lockhart, Robert S. (1972): Levels of processing: A framework for memory research. In: *Journal of Verbal Learning and Verbal Behavior* 11: 6, 671-684.
- Cunningham, D. Joseph (2019): L2 Pragmatics Learning in Computer-Mediated Communication. In: Taguchi, Naoko (Ed.): *The Routledge Handbook of Second Language Acquisition and Pragmatics*. New York: Routledge, 372-386.

- Cutler, Cecilia (1999): Yorkville Crossing: White teens, hip hop and African American English. In: *Journal of Sociolinguistics* 3: 4, 428-442.
- Dannerer, Monika (2014): Sprachbiographische Äußerungen und Erzählerweb im Längsschnitt als Zugangswege zur Beschreibung von Zweitspracherwerb. In: Ahrenholz, Bernt / Grommes, Patrick (Eds.): *Zweitspracherwerb im Jugendalter*. Berlin / Boston: de Gruyter, 295-317.
- Deshors, Sandra C. / Gries, Stefan Th. (2023): Using corpora research on second language psycholinguistics. In: Godfroid, Aline / Hopp, Holger (Eds.): *The Routledge handbook of second language acquisition and psycholinguistics*. London: Routledge, 164-177.
- Dewaele, Jean-Marc (2013): Learner-internal psychological factors. In: Herschensohn, Julia / Young-Scholten, Martha (Eds.): *The Cambridge handbook of second language acquisition*. 1st ed. Cambridge: Cambridge University Press, 159-179.
- Dizon, Gilbert (2022): YouTube for second language learning: What does the research tell us? In: *Australian Journal of Applied Linguistics* 5: 1, 19-26.
- Dörnyei, Zoltán (2001): New themes and approaches in second language motivation research. In: *Annual Review of Applied Linguistics* 21, 43-59.
- Dörnyei, Zoltán (2003): Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. In: *Language learning* 53, 3-32.
- Dörnyei, Zoltán (2009): Individual Differences: Interplay of Learner Characteristics and Learning Environment. In: *Language Learning* 59, 230-248.
- Dunn, Karen / Iwaniec, Janina (2022): Exploring the relationship between second language learning motivation and proficiency: A latent profiling approach. In: *Studies in Second Language Acquisition* 44: 4, 967-997.
- Erbaggio, Pierluigi / Gopalakrishnan, Sangeetha / Hobbs, Sandra / Liu, Haiyong (2012): Enhancing Student Engagement through Online Authentic Materials. In: *IALLT Journal of Language Learning Technologies* 42: 2, 27-51.
- Fernández, Julieta / Davis, Tracy S. (2021): Overview of Available Learner Corpora. In: Paquot, Magali / Tracy-Ventura, Nicole (Eds.): *The Routledge handbook of second language acquisition and corpora*. London: Routledge, 145-157.
- Friginal, Eric / Roberts, Jennifer (2022): Corpora for Materials Design. In: Jablonkai, Reka R. / Csomay, Eniko (Eds.): *The Routledge Handbook of Corpora and English Language Teaching and Learning*. London: Routledge, 131-146.
- García, Ofelia / Leiva, Camila (2014): Theorizing and Enacting Translanguaging for Social Justice. In: Blackledge, Adrian / Creese, Angela (Eds.): *Heteroglossia as Practice and Pedagogy*. Dordrecht: Springer, 199-216.
- Gilmore, Alex (2007): Authentic materials and authenticity in foreign language learning. In: *Language Teaching* 40: 2, 97-118.
- Gilquin, Gaëtanelle / Granger, Sylviane (2022): Using data-driven learning in language teaching. In: O'Keefe, Anne / McCarthy, Michael (Eds.): *The Routledge Handbook of Corpus Linguistics*. 2nd Edition. London: Routledge, 430-442.

- Glaznieks, Aivars / Frey, Jennifer-Carmen (2020): Das DiDi-Korpus: Internetbasierte Kommunikation aus Südtirol. In: Marx, Konstanze / Lobin, Henning / Schmidt, Axel (Eds.): *Deutsch in Sozialen Medien: Interaktiv – multimodal – vielfältig*. Berlin / Boston: de Gruyter, 353-354.
- Goschler, Juliana / Stefanowitsch, Anatol (2014): Korpora in der Zweitspracherwerbsforschung: Sieben Probleme aus korpuslinguistischer Sicht. In: Ahrenholz, Bernt / Grommes, Patrick (Eds.): *Zweitspracherwerb im Jugendalter*. Berlin / Boston: de Gruyter, 341-360.
- Granger, Sylviane (2009): The contribution of learner corpora to second language acquisition and foreign language teaching: A critical evaluation. In: Aijmer, Karin (Ed.): *Corpora and language teaching*. Amsterdam: John Benjamins, 13-32.
- Gysin, Daniel (2014): ALTER marias bild im svz hat schon style:-D – Öffentliche und nicht-öffentliche Kommunikation Jugendlicher in sozialen Netzwerken. In: Kotthoff, Helga / Mertzlufft, Christine (Eds.): *Jugendsprachen: Stilisierungen, Identitäten, mediale Ressourcen*. Frankfurt am Main: Peter Lang, 215-244.
- Harklau, Linda (2007): The Adolescent English Language Learner. In: Cummins, Jim / Davison, Chris (Eds.): *International Handbook of English Language Teaching*. Boston: Springer, 639-653. https://doi.org/10.1007/978-0-387-46301-8_41.
- Hee, Katrin (2018): Jugendkommunikation in schulischen Lehr-/Lernkontexten: Haupt- und Nebenkommunikation im Vergleich. In: Ziegler, Arne (Ed.): *Jugendsprachen: Aktuelle Perspektiven Internationaler Forschung*. Berlin / Boston: de Gruyter, 269-302.
- Höhn, Sviatlana (2015): Corpus of long-term instant messaging based dialogues between advanced learners of German as a foreign language and German native speakers: deL1L2IM. ELRA. <https://orbilu.uni.lu/handle/10993/20579>.
- Horwitz, Elaine K. / Horwitz, Michael B. / Cope, Joann (1986): Foreign language classroom anxiety. In: *The Modern language journal* 70: 2, 125-132.
- James, Carl (2013): *Errors in Language Learning and Use: Exploring Error Analysis*. London: Routledge.
- Jauregi-Ondarra, Kristi / Canto, Silvia / Melchor-Couto, Sabela (2022): Virtual Worlds and Second Language Acquisition. In: Ziegler, Nicole / González-Lloret, Marta (Eds.): *The Routledge handbook of second language acquisition and technology*. New York: Routledge, 311-325.
- Kabooha, Raniah / Elyas, Tariq (2018): The Effects of YouTube in Multimedia Instruction for Vocabulary Learning: Perceptions of EFL Students and Teachers. In: *English Language Teaching* 11: 2, 72-81.
- Kehily, Mary Jane / Nayak, Anoop (1997): 'Lads and Laughter': Humour and the production of heterosexual hierarchies. In: *Gender and Education* 9: 1, 69-88.
- Keim, Inken (2007): *Die 'türkischen Powergirls': Lebenswelt und kommunikativer Stil einer Migrantinnengruppe in Mannheim*. Tübingen: Narr.
- Kerswill, Paul (2013): Identity, ethnicity and place: the construction of youth language in London. In: Auer, Peter / Hilpert, Martin / Stukenbrock, Anja / Szmrecsanyi, Benedikt (Eds.): *Space in language and linguistics*. Berlin / Boston: de Gruyter, 128-164.
- Koch, Peter / Österreicher, Wulf (2007): Schriftlichkeit und kommunikative Distanz. In: *Zeitschrift für germanistische Linguistik* 35: 3, 346-375. <https://doi.org/10.1515/zgl.2007.024>.

- Krashen, Stephen (1982): *Principles and practice in second language acquisition*. Oxford: Pergamon Press.
- Kruk, Mariusz (2019): Dynamicity of perceived willingness to communicate, motivation, boredom and anxiety in Second Life: the case of two advanced learners of English. In: *Computer Assisted Language Learning* 35: 1-2, 190-216.
- Krumm, Hans-Jürgen / Jenkins, Eva-Maria (2001): *Kinder und ihre Sprachen - lebendige Mehrsprachigkeit: Sprachenportraits gesammelt und kommentiert von Hans-Jürgen Krumm*. Wien: Eviva.
- Lee, Ju Seong / Lu, Ying (2021): L2 motivational self system and willingness to communicate in the classroom and extramural digital contexts. In: *Computer Assisted Language Learning* 36: 1-2, 126-148.
- Leibniz-Institut für Deutsche Sprache (2022): IDS: Korpuslinguistik: Korpusausbau. <http://www1.ids-mannheim.de/kl/projekte/korpora.html> (21.07.2023).
- Li, Li (2017): *New Technologies and Language Learning*. London: Palgrave Macmillan.
- Lorenzo-Dus, Nuria / Blitvich, Pilar Garcés-Conejos / Bou-Franch, Patricia (2011): On-line polylogues and impoliteness: The case of postings sent in response to the Obama Reggaeton YouTube video. In: *Journal of Pragmatics* 43: 10, 2578-2593.
- Lüngen, Harald / Kupietz, Marc (2020): IBK- und Social Media-Korpora am Leibniz-Institut für Deutsche Sprache. In: Marx, Konstanze / Lobin, Henning / Schmidt, Axel (Eds.): *Deutsch in Sozialen Medien: Interaktiv – multimodal – vielfältig*. Berlin / Boston: de Gruyter, 319-342.
- MacIntyre, Peter D. / Clément, Richard / Dörnyei, Zoltán / Noels, Kimberly A. (1998): Conceptualizing Willingness to Communicate in a L2: A Situational Model of L2 Confidence and Affiliation. In: *The Modern Language Journal* 82: 4, 545-562.
- MacIntyre, Peter D. (2002): Motivation, anxiety and emotion in second language acquisition. In: Robinson, Peter (Ed.): *Individual Differences and Instructed Language Learning*. Amsterdam: John Benjamins, 45-68.
- MacIntyre, Peter D. / Baker, Susan C. / Clément, Richard / Donovan, Leslie A. (2003): Sex and age effects on willingness to communicate, anxiety, perceived competence, and L2 motivation among junior high school French immersion students. In: *Language learning* 53, 137-166.
- MacIntyre, Peter D. (2007): Willingness to Communicate in the Second Language: Understanding the Decision to Speak as a Volitional Process. In: *The Modern Language Journal* 91: 4, 564-576.
- Marchand, Tim / Akutsu, Sumie (2015): First steps in assigning proficiency to texts in a learner corpus of computer-mediated communication. In: Callies, Marcus / Götz, Sandra (Eds.): *Learner corpora in language testing and assessment*. Amsterdam: John Benjamins, 85-112.
- Marcoccia, Michel (2004): On-line polylogues: conversation structure and participation framework in internet newsgroups. In: *Journal of Pragmatics* 36: 1, 115-145.
- Meskill, Carla / Anthony, Natasha (2005): Foreign language learning with CMC: forms of online instructional discourse in a hybrid Russian class. In: *System* 33: 1, 89-105.
- Michaelides, Michalis P. / Brown, Gavin T. L. / Eklöf, Hanna / Papanastasiou Elena C. (2019): *Motivational Profiles in TIMSS Mathematics: Exploring Student Clusters Across Countries and Time*. Cham: Springer International Publishing.

- Milroy, James / Milroy, Lesley (1999): *Authority in Language: Investigating Standard English*. 3rd ed. London: Routledge.
- Mitchell, Rosamund (2020): Corpora and Instructed Second Language Acquisition. In: Paquot, Magali / Tracy-Ventura, Nicole (Eds.): *The Routledge handbook of second language acquisition and corpora*. London: Routledge, 252-264.
- Mizumoto, Tomoya et al. (2011): Mining Revision Log of Language Learning SNS for Automated Japanese Error Correction of Second Language Learners. In: *Proceedings of 5th International Joint Conference on Natural Language Processing*. Chiang Mai, Thailand: Asian Federation of Natural Language Processing, 147-155. <https://aclanthology.org/I11-1017>.
- Myhill, Debra / Newman, Ruth (2016): Metatalk: Enabling metalinguistic discussion about writing. In: *International Journal of Educational Research* 80, 177-187.
- Myhre, Tone Stuler / Fiskum, Tove Anita (2021): Norwegian teenagers' experiences of developing second language fluency in an outdoor context. In: *Journal of Adventure Education and Outdoor Learning* 21, 201-216.
- Nowotny, Andrea (2005): Daumenbotschaften. Die Bedeutung von Handy und SMS für Jugendliche. In: *Net-Worx*. <http://www.mediensprache.net/networx/networx-44.pdf> (21.07.2023).
- Palacios Martínez, Ignacio (2018): Lexical Innovation in the Language of Teenagers. A Cross-Linguistic Perspective. In: Ziegler, Arne (Ed.): *Jugendsprachen: Aktuelle Perspektiven Internationaler Forschung*. Berlin / Boston: de Gruyter, 363-390.
- Paquot, Magali (2022): Corpora and Second Language Acquisition. In: Jablonkai, Reka R. / Csomay, Eniko (Eds.): *The Routledge Handbook of Corpora and English Language Teaching and Learning*. London: Routledge, 26-40.
- Peterson, Mark (2016): The use of massively multiplayer online role-playing games in CALL: an analysis of research. In: *Computer Assisted Language Learning* 29: 7, 1181-1194.
- Prado, Malila C.A. / Tosquil Lucks, Patricia (2019): Designing the Radiotelephony Plain English Corpus (RTPEC): A specialized spoken English language corpus towards a description of aeronautical communications in non-routine situations. In: *Research in Corpus Linguistics* 7, 113-128.
- Reinders, Hayo / Hubbard, Philip (2013): CALL and learner autonomy: Affordances and constraints. In: Thomas, Michael / Reinders, Hayo / Warschauer, Mark (Eds.): *Contemporary Computer-Assisted Language Learning*. Bloomsbury Academic, 359-375.
- Reinders, Hayo / Wattana, Sorada (2014): Can I Say Something? The Effects of Digital Gameplay on Willingness to Communicate. In: *Language Learning & Technology* 18: 2, 101-123.
- Ritchie, Mathy / Black, Catherine (2012): Public Internet Forums: Can They Enhance Argumentative Writing Skills of Second Language Learners? In: *Foreign Language Annals* 45: 3, 349-361.
- Russell, Victoria (2020): Language anxiety and the online learner. In: *Foreign Language Annals* 53: 2, 338-352.
- Sachs, Rebecca / Baralt, Melissa / Gurzynski-Weiss, Laura (2023): Psycholinguistics for the language classroom. In: Godfroid, Aline / Hopp, Holger (Eds.): *The Routledge handbook of second language acquisition and psycholinguistics*. London: Routledge, 387-399.

- Saferinternet.at (2018): Jugend-Internet-Monitor 2018. In: *Saferinternet.at*. <https://www.saferinternet.at/presse-detail/jugend-internet-monitor-2018/> (21.07.2023).
- Silverstein, Michael (2003): Indexical order and the dialectics of sociolinguistic life. In: *Language & communication* 23: 3-4, 193-229.
- Smith, Bryan (2005): The Relationship between Negotiated Interaction, Learner Uptake, and Lexical Acquisition in Task-Based Computer-Mediated Communication. In: *TESOL Quarterly* 39: 1, 33-58.
- Statista (2023a): YouTube – Anteil der Nutzer nach Altersgruppen in Deutschland 2022. In: *Statista*. <https://de.statista.com/statistik/daten/studie/511076/umfrage/anteil-der-nutzer-von-youtube-nach-altersgruppen-in-deutschland/> (21.07.2023).
- Statista (2023b): Instagram – Nutzerstruktur nach Altersgruppen in Deutschland 2022. In: *Statista*. <https://de.statista.com/statistik/daten/studie/771453/umfrage/nutzerstruktur-von-instagram-nach-altersgruppen-in-deutschland/> (21.07.2023).
- Taylor, Florentina (2013): *Self and Identity in Adolescent Foreign Language Learning*. Bristol: Multilingual Matters.
- Terantino, Joe (2011): YouTube for foreign languages: You have to see this video. In: *Language Learning & Technology* 15: 1, 10-16.
- Thompson, Dominic et al. (2016): Emotional responses to irony and emoticons in written language: Evidence from EDA and facial EMG. In: *Psychophysiology* 53: 7, 1054-1062.
- Tomlinson, Brian (2012): Materials development for language learning and teaching. In: *Language Teaching* 45: 2, 143-179.
- Ushioda, Ema / Dörnyei, Zoltán (2012): Teaching and researching motivation. In: Gass, Susan M. / Mackey, Alison (Eds.): *The Routledge handbook of second language acquisition*. London: Routledge, 396-409.
- Villoria-Prieto, Javier / Suso López, Javier (2018): The Use of Dialogues in Teaching Foreign Languages (Sixteenth Century): Circulations and Adaptation of Berlaimont's *Dictionarium* (1556) in Spain, the Netherlands, and England. In: McLelland, Nicola / Smith, Richard (Eds.): *The History of Language Learning and Teaching*. Cambridge: Legenda, 67-82.
- Vyatkina, Nina / Belz, Julie A. (2006): A learner corpus-driven intervention for the development of L2 pragmatic competence. In: Bardovi-Harlig, Kathleen / Félix-Brasdefer, J. César / Omar, Alwiya S. (Eds.): *Pragmatics and Language Learning*. Honolulu: National Foreign Language Resource Center, 315-357.
- Wulff, Stefanie (2020): Usage-based Approaches. In: Paquot, Magali / Tracy-Ventura, Nicole (Eds.): *The Routledge handbook of second language acquisition and corpora*. London: Routledge, 175-188.

Biographical note: Louis Cotgrove is a researcher at the Leibniz Institute for the German Language in Mannheim, and specialises in corpus linguistic methods to analyse youth and digital linguistic phenomena, such as intensifiers, emoji, and language change. He can also be found programming APIs for lexical resources as part of the Germany-wide Text+ project (<https://www.text-plus.org/>).

Contact address:

Dr. Louis Cotgrove
Leibniz-Institut für Deutsche Sprachen
68161, Mannheim
Germany
cotgrove@ids-mannheim.de



Lizenz: CC BY 4.0 International - Creative Commons, Namensnennung.