

LEARNER CORPUS RESEARCH MEETS FOREIGN LANGUAGE EDUCATION: EXAMPLES FROM THE SWISS LEARNER CORPUS SWIKO

Korpuslinguistische Anwendungen in pädagogischen Kontexten sind bisher eher Wunsch als Wirklichkeit. Dieser Beitrag stellt das mehrsprachige Schweizer Lernerkorpus SWIKO vor und zeigt anhand von zwei Szenarien, wie dieses Korpus beim Sprachenlehren und -lernen genutzt werden kann. SWIKO umfasst schriftliche und mündliche task-basierte Texte von SchülerInnen der Sekundarstufe I in den Schulfremdsprachen (Deutsch, Englisch und Französisch) und in der jeweiligen Unterrichtssprache. Szenario 1 argumentiert anhand von Analysen der fremdsprachlichen Texte, dass Lernerleistungen bezogen auf Aufgaben-Typen illustriert werden sollten. Szenario 2 nutzt u.a. das Subkorpus der Unterrichtssprache, um authentische Lernmaterialien zu entwickeln. Als Veranschaulichung dienen Arbeitsblätter zur Negation im Deutschen.

● Nina Hicks
| Université de Fribourg
Thomas Studer
| Université de Fribourg

Introduction

According to Römer (2008; see also e.g., Flinz, 2021), there are two ways in which corpora – large collections of oral or written texts (Lemnitzer & Zinsmeister, 2015; cf. introduction of this edition) – can be applied in pedagogical settings: either indirectly via researchers and material writers; or directly via learners and teachers in the classroom.

In the indirect form, corpora and corpus linguistic findings contribute to the development of reference works, teaching materials, and teacher training. The authentic and typical examples of actual language use inform decisions on what should be taught and in what order (McEnergy & Xiao, 2011). Learner corpora have increasingly contributed to this process, offering valuable insights on the mechanisms and challenges of language learning, for example based on analyses of errors or contrastive over- and underuse (for an overview see e.g., Meunier, 2020). More recently, the influ-

ence of other variables such as tasks on differences in language use has attracted researchers' attention (e.g., Alexopoulou et al., 2017)

In the direct form, referred to as *data-driven learning* (DDL), teachers and learners use corpus linguistic techniques and tools for pedagogical purposes. In the last decades, corpus linguists have highlighted the advantages of the DDL approach, particularly the exposure to rich, authentic language use as well as the autonomous and collaborative discovery approach (e.g., Gilquin & Granger, 2010; McEnergy & Xiao, 2011). In line with the slogan "Every student a Sherlock Holmes" (Johns, 1997, p. 101), students are encouraged to observe enhanced corpus data, and, based on what they notice, hypothesize, generate rules, and check their insights into linguistic patterns. This active and autonomous involvement also allegedly increases motivation.

Despite these proclaimed advantages, the DDL approach has yet to find its way

into the everyday classroom (Meunier, 2020). Several factors contribute to this scarce uptake (e.g., Meunier, 2019; Römer, 2009; Tribble, 2008): Technical tools have to be available and, more importantly, mastered – a time-consuming endeavour for both teachers and learners, especially considering that the topic and approach have to be carefully selected and adapted to the specific context. Furthermore, while research results have shown positive effects of DDL on learning gains (e.g., Boulton & Cobb, 2017), most studies were conducted by corpus linguists (as opposed to teachers) with advanced and tertiary learners focusing on English as a foreign language (although see recent work by Crosthwaite, 2020 for younger learners; or Vyatkina, 2020 for German as a foreign language).

With these gaps in mind, the next section provides a brief overview of the SWIKO project and corpus, followed by two scenarios of how the corpus can be utilized in two educational settings: teacher training and the secondary school classroom.

SWIKO

SWIKO is a multilingual corpus currently being developed at the Institute of Multilingualism in Fribourg (CH). Following the trend towards communication and, more recently, action-orientation in foreign language education in Switzerland (e.g., Bertschy et al., 2015; Studer, 2023), the project investigates what vocabulary and grammar look like under these leading principles in learner texts at the end of mandatory schooling.

Between 2017 and 2022, data was collected among 14- to 17-year old students attending grades 10-12 (HarmoS) in both German- and French-speaking Switzerland as well as a German-English bilingual school in Eastern Switzerland. Students completed eight communicative tasks which were systematically varied by intended text type, topic, and task structure (Ellis et al., 2020). The corpus currently contains over 3'000 annotated written and spoken, paper- and computer-based productions by Swiss lower secondary school students in three languages (German, French and English), both as their language of schooling and foreign languages.

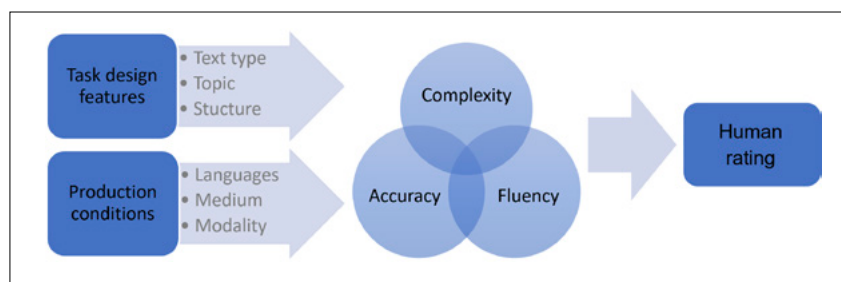


Figure 1
Overview of variables and analysis workflow in the SWIKO project.

The productions were then processed and analysed using corpus-linguistic methods. First, additional linguistic information was annotated (Schmid, 2013), such as the lemma (e.g., the tokens *gehe*, *gehst* und *ging* all stem from the lemma *gehen*) and part-of-speech (e.g., whether it's a verb or a noun). This allows us to analyse how long, diverse, dense, and sophisticated each learner production is (based on the CAF framework, Housen et al., 2012). Additionally, a *target hypothesis* – a grammatically and orthographically correct version (Lüdeling & Hirschmann, 2015) – was formulated for each learner text. By comparing the two versions, we can categorize the types of errors and, in extension, identify particularly challenging structures for learners. Finally, prospective foreign language teachers rated the productions based on the scales of the Common European Framework of Reference (Council of Europe, 2001, 2020). This enables us to examine which task design features and linguistic aspects correlate with human ratings of the learner texts (Figure 1).

In sum, grounded in Granger's (2015) Contrastive Interlanguage Analysis, we have been investigating how different tasks affect the linguistic properties of the resulting productions and how these relate to the human ratings (e.g., Karges et al., 2019, 2022; Studer & Hicks, 2022; Weiss et al., 2022). Based on our findings, the following chapters present two scenarios of how the SWIKO corpus can be used in educational settings. The first highlights some findings on how task design features affect the linguistic properties of the learner productions. The second discusses possibilities on how to use concordances from the German sub-corpus to develop teaching material for the secondary school classroom.



Nina Hicks works as a Research Assistant for the SWIKO/WETLAND project at the Institute of Multilingualism

Fribourg. She holds an MA in Foreign Language Education from the University of Fribourg and has extensive experience teaching primary school students and German as a foreign language learners in Switzerland, California, and the UK.



Thomas Studer is a full Professor of German as a foreign and second language at the Department of Multilingualism

and Foreign Language Education at the University of Fribourg and member of the board of directors of the Institute of Multilingualism in Fribourg. His research interests include corpus-based language acquisition, foreign language education, and language testing and assessment.

Showcasing task-based variation of learner language

To reflect the scope of language use that students encounter in their instructed foreign language classes, the SWIKO team developed eight different tasks (Ellis et al., 2020). As mentioned earlier, the task design features were systematically varied by intended text type, topic, and structure. As a result, students described and argued on more personal and more academic topics based on more and less restrictive input (e.g., answering short personal questions as opposed to writing a self-portrait without specific guidelines). Table 1 provides an overview of the tasks and variables.

ID	Description of task	Task type		Topic		Structure	
		des	arg	acad	pers	more	less
SWI01	Answer short personal questions	x		x		x	
SWI02	Describe graphs about pets in Switzerland	x			x	x	
SWI03	Discuss a list of vacation options		x	x		x	
SWI04	Discuss a list of the most important inventions		x		x	x	
SWI05	Create a self-portrait for a class exchange	x		x			x
SWI06	Present a topic (8 options)	x			x		x
SWI07	Discuss a later school start and finish		x	x			x
SWI08	Discuss replacement of foreign language classes with language exchange abroad		x		x		x

Table 1

Tasks and task variation in the SWIKO corpus. Task type was characterized as descriptive or argumentative, topic as academic or personal, and structure as more or less restrictive input given in the task prompt.

The systematic variation of task design features allows for insights on how the type of task affects the production. We analysed the linguistic aspects of the productions based on the CAF framework, which distinguishes three main components: “*complexity* is commonly characterized as the ability to use a wide and varied range of sophisticated structures and vocabulary in the L2, *accuracy* as the ability to produce target-like and error-free language, and *fluency* as the ability to produce the L2 with native-like¹ rapidity, pausing, hesitation, or reformulation” (Housen et al., 2012, p. 2, emphasis added).

In the next paragraphs, we present selected CAF differences based on two task design features – text type and topic – among 544 written German as a foreign language (DaF) productions in the SWIKO corpus. We believe that such insights can provide an illustrative data base for teacher training: On the one hand, it

offers near-authentic examples of learner language, which can help trainees in developing a more realistic and nuanced understanding of students’ output at the end of mandatory schooling. On the other hand, it can raise awareness of the importance of carefully selecting tasks for training and assessments as the type of task heavily affects the extent to which learners can expand on and demonstrate their linguistic competence.

Complexity

Among the written DaF productions in the SWIKO corpus, we observed that particularly the intended text type played an important role in regards to linguistic complexity: descriptive texts contained denser and more sophisticated vocabulary as well as more coordinated but fewer dependent clauses per sentence. In other words, students used more lexical words (particularly nouns) in their descriptions (mean noun ratio 0,32 vs. 0,18), which are often not found among the most frequent words of a language (mean logarithmic token frequency 4,63 vs. 4,31 based on SUBTLEX-DE, similar results based on Google Books 2000, Open Subtitles and DeReWo corpora). Furthermore, students were twice as inclined to link two or more main clauses with the co-ordinating conjunction *und* in their descriptive texts (mean coordinate phrases per clause 0,27 vs. 0,12). In contrast, when writing argumentative texts, they wrote longer sentences (mean sentence length in tokens 11,51 vs. 8,99) and used sub-ordinating conjunctions such as *weil*, *dass* or *wenn* three times as often to connect a main and a dependent clause (mean dependent clause ratio 0,32 vs. 0,11).

We exemplify these differences with two tasks which asked students to produce texts on an academic topic with more restrictive input, but which differ in terms of intended text type: a description of a graph about pets in Switzerland (SWI02) and a discussion of the allegedly most important inventions (SWI04). Both of the following texts were written by the same student (Figure 2).

The description of the graph is lexically denser and more sophisticated, as it contains more nouns (noun ratio 0,24 vs. 0,13) and fewer common words (mean SUBTLEX token log frequency 4,70 vs. 4,30). Conversely, the grammatical complexity is higher in the argumentation

¹ Although the term *nativeness* has been debated, with scholars highlighting its vagueness (e.g., Cheng et al., 2021) and the ignorance of intra-individual variation among L1 groups (Shadrova et al., 2021). In SWIKO, we described our comparison group as “*language of schooling*”.

as sentences are longer (26 vs. 14 tokens per sentence), as well as more varied and complex. Specifically, the argumentative text contains five main clauses linked with the co-ordinating conjunctions *und* and *aber*, as well as four dependent clauses linked with the sub-ordinating conjunctions *dass* and *weil*. In comparison, the description consists of six main clauses exclusively linked with *und* and one dependent clause linked with *dass*.

Accuracy and Fluency

In order to investigate the accuracy of the written German productions of the SWIKO corpus in more details, a *target hypothesis* was formulated for each learner text. Comparing the two versions allows us to accurately trace the types of learner errors and develop a more refined picture of which structures proved particularly challenging for the learners. We measured accuracy in three aspects (Table 2): 1) the ratio of words written in a language other than German (“non-target” words); 2) the ratio of orthographic errors including capitalization, graphemes, and word boundary; and 3) the ratio of grammatical errors. This last aspect is again divided into four sub-categories: a) missing words, b) unnecessarily added words, and c) erroneously chosen or inflected words, all of which are measured at token level, as well as d) wrongly positioned constituents, which are measured as the ratio of (in)correct sentences. Table 2 provides an overview of the categories with examples.

In written productions, fluency is often equated with the length of the text, operationalized as the number of tokens² (Wolfe-Quintero et al., 1998).

Among written DaF productions in the SWIKO corpus, at the token level, the topic influenced both fluency and accuracy: in contrast to more academic topics, tasks on more personal topics generally resulted in longer (number of tokens 47 vs. 39) and more accurate texts (ratio of correct tokens 0,66 vs. 0,55). More specifically, they contained fewer errors both at the orthographic (error ratio 0,12 vs. 0,14) and grammatical level (error ratio 0,15 vs. 0,22), and more tokens in the target language (ratio of non-target tokens 0,06 vs. 0,10).

For example, when comparing two descriptive and less structured tasks,

SWI02: Graph about pets (descriptive) Student Ri211, 55 tokens	SWI04: Important inventions (argumentative) Student Ri211, 52 tokens	
<i>In diese graphique du hast 4 punkte[.] die ersten is über Wo hast ein tier und wir denken das eins out of zwei perssonen hast ein tier[.] punkte 2 wir haben 8.2 millionen tiere in der schweis[.] punkte 3 eine Katze catch zu viele other tiere und punkte 4 wir spend 800 millionen chf pro jahre fur die tiere in schweis[.]</i>	<i>ich denke das diese inventionen sind in die liste weil das sehr wichtig sind und die ordeung ist gut fur die fünf ersten aber icht fur die fünf lesten[.] ich denke das die car ist mehr wichtig als die brille und die uhr must gehen zent placen weil das sehr wichtig ist[.]</i>	
0,24	noun ratio	0,13
4,30	mean SUBTLEX token log frequency	4,70
14	tokens per sentence	26
6 (<i>und</i>)	main clauses (<i>coordinating conjunctions</i>)	5 (<i>und, aber</i>)
1 (<i>dass</i>)	dependent clauses (<i>subordinating conjunctions</i>)	4 (<i>dass, weil</i>)

Figure 2

A descriptive and an argumentative text by the same student.

Category	Sub-category	Examples
Non-target	Language of schooling	<i>papillon (Schmetterling), ajutiere (hinzufügen)</i>
	Other languages	<i>glasses (Brille), o'clock (Uhr)</i>
Orthography	Capitalization	<i>computer, Acht, LieblingsTiere</i>
	Graphemes	<i>Personnen, Tire, bezucht, obwohl, helfbereit</i>
	Word boundary	<i>Lieblings Musik, Diewochenende, Liebst-du</i>
Grammar	Missing word	<i>In [der] Schweiz</i>
	Unnecessary word	<i>Es gibt 8,2 Mio. [die] Tiere</i>
	Wrongly inflected	<i>8,2 Million[en], du [hat], mit [deinem] Freunden</i>
	Wrongly chosen	<i>ich [habe] 14 Jahre alt, Ferien [im] Meer</i>
	Wrong position	<i>In der Schweiz [es gibt] viele Tiere.</i>

Table 2

Error categories with examples as analysed among the German productions in the SWIKO corpus.

² Our token count excludes punctuation and numerals, and each entity is counted as one token irrespective of the number of words, e.g., *Titanic*, *Harry Potter and the Philosopher's Stone*, or *Nintendo Switch* are each considered a one entity-token.

students wrote longer and more accurate texts presenting themselves (SWI05, more personal) as opposed to presenting a topic such as languages or oceans (SWI06, more academic). Again, the following two texts were written by the same student (Figure 3). The more personal text is longer (67 vs. 47 tokens) and contains fewer non-target words (ratio 0,02 vs. 0,17), fewer orthographic errors (ratio 0,06 vs. 0,15) and fewer grammatical errors (ratio 0,12 vs. 0,28).

SWI05: Anonymized self-portrait (personal) Student Ri300, 67 tokens		SWI04: Topic presentation (academic) Student Ri300, 47 tokens	
<p>Hallo! Ich heisse Laura, ich bin 14 Jahre alt und wohne in Boudry. Ich habe ein grosse Bruder, ich habe keine Haustiere[,] aber ich mochte gern ein Hund haben. Ich spiele Tennis und ich mag gern shoppen. Ich bin sehr freundlich und helfbereit. Meine Lieblingsfahrt ist Mathe und ich liebe gar nicht Französisch. Meine Lieblingsfarbe ist blue. Meine Lieblingsfilme ist „Titanic“ und meine Liebingsserie ist „Teen Wolf“. Bis bald!</p>		<p>Es ist 6 000 zu 7 000 Langue am Erde. Wir spreche am mehrest Langue in Asie und Afrique. Es ist viel Sweisprachig und Dreisprachig Leute. [Die] Leute spricht am mehresten English und Spanish und am meisten Deutsch und Französisch, in Europe. In der Schweiz [die] Leute spricht Deutsch, Französich, Italienish und Romanch.</p>	
2 %	non-target token (pink)	17 %	
6 %	orthographic error (blue)	15 %	
12 %	grammatical error (yellow)	26 %	

Figure 3

A text on a personal and a text on an academic topic by the same student.

Conversely, at the sentence level, the text type mattered most: sentences in descriptive texts were more often written in the correct order than sentences in argumentative texts (mean ratio of correct sentences 0,73 vs. 0,47; Figure 3 ratios 1,00 vs. 0,80). This could reflect the types of structures used: as reported in the complexity section, students wrote more subordinate phrases in argumentative productions, which require the more difficult verb final position and are therefore more prone to errors.

Negation in the DaF classroom

During our accuracy analysis (see section above), negation was revealed as a particularly challenging structure: German as a foreign language learners used *nicht* instead of *kein* in 54% of cases where *kein* was required, compared to just 2% among their German as the language of schooling peers. Moreover, it seems that learners apply the “*kein* is followed by a noun” rule correctly once they know it as it was never used erroneously; neither instead of *nicht* nor in the wrong position³.

The following paragraphs therefore present ideas on how the German sub-corpora can be used in the German as a foreign language secondary school classroom to learn about negation. However, the underlying mechanisms of the task creation could be adapted to other topics such as capitalization or sub- and co-ordinating conjunctions (see Vyatkina, 2020 for examples on how to use the DWDS corpus in the DaF classroom).

DDL activities can be categorized on a continuum from teacher- to learner-led, and from relatively controlled such as gap exercises to open-ended, although more support might be beneficial for younger and less proficient learners (Gilquin & Granger, 2010). While inductive approaches are usually favoured in DDL, we aim to offer a wide spectrum of tasks and exercises as well as suggestions on alternatives for differentiation so that teachers can decide on whichever option best suits their class. Furthermore, our material (Übungsblatt Negation 1-5) is built around concordance lines extracted from SWIKOweb⁴, i.e., lines of texts taken from the corpus and displayed with the highlighted key word in the middle. Other corpus data such as frequency lists or word clouds can be generated and used, for example to brainstorm vocabulary before a writing assignment (see Figure 4 for examples).

German as the language of schooling corpus

Concordances from the German as the language of schooling sub-corpus can be used in a variety of ways. As an introduction, these concordances can serve as an illustration from which learners derive rules – for example, *kein* is always followed by a noun, whereas *nicht* is often used in conjunction with adjectives or adverbs. We designed two difficulty levels, both of which follow an illustration – interaction – induction approach (Wicher, 2019) in which learners are encouraged to cooperate with their class peers. In an easier version, students derive rules from lists which were pre-sorted by the teacher according to the two rules (Übungsblatt Negation 1a). A more challenging version prompts students to group individual concordance lines by type and then derive the rules based on their observations (Übungsblatt

3 There were a few inflection errors (18%), e.g., *kein* instead of *keinen*. However, these only occurred slightly more frequently than among their peers with German as the language of schooling (8%).

4 Currently accessible via a personalized login.

Negation 1b). In either version, once the rules have been established, students can add further concordance lines from the corpus or their own examples.

In order to consolidate their knowledge, the students then create a mind map with collocations or chunks in which each of the two negation words commonly appear (Übungsblatt Negation 2). For scaffolding, the concordance lines from the introduction can be re-used (Übungsblatt Negation 1b), or additional concordance lines printed, or, if there is enough time, students could look up concordances themselves. Alternatively, a “class mind map” can be continuously and collectively expanded throughout the unit. In a following step, students complete gap exercises (Übungsblatt Negation 3), where they have to decide on which negation word is appropriate. This type of exercise can be created easily by omitting the highlighted *key word* of the concordance lines. A second exercise asks them to combine words from a list to common chunks, which was collected from corpus examples.

German as a foreign language corpus

Alternatively, concordances from the German as a foreign language sub-corpus can be used (Übungsblatt Negation 4): As a starting point, students reflect on how negation is formed in their language of schooling and any other languages they might know, before discussing their observations with their classmates and comparing them to the formation of negation in German. Then, they are asked to correct excerpts from the SWIKO corpus written by their German as a foreign language learning peers. In the exercise offered, students first have to decide whether a given concordance line actually does contain an error, and only correct it if necessary. Again, the level of difficulty can be adapted: In an easier version, a few concordance lines which all contain the same type of error can be offered. In a more challenging version, the students can additionally be asked to correct other types of errors in the concordance lines, such as orthographic or other types of grammatical errors. In either scenario, after investigating learner problems, it is recommended to highlight appropriate or correct use in follow-up exercises (Granger & Tribble, 1998).

ein Kuh in ein Jahre.Ein Foyer uns zwei haben ein **tier**. In Schweiz barcht millions zweihundert Tiere die Frank Schweiz in JahreEin Familie in zwei hat ein **Tier**. Ein Katze, fur ein Monate hat 400 Mouse. Das ist ur ein Monate hat 400 Mouse. Das ist 8.2 millions **Tier** domestik in Schweiz. Die dépense ist 1500 CHF fur dein Hund und 1000 fur Jahrein Katze.1/2 hat ein **Tier** in die Hause. 2. 8,2 Millio d'animaux. 3. Eine Ka 3'500 fr eine Kuh.Eine personen sur zehn hat ein **Tier**. 8.2000000 Haustier.Eine Familie fur 2 habst ein e du hast 4 punkte die ersten is über Wo hast ein **tier** und wir denken das eins out of zwei perssonen has ir denken das eins out of zwei perssonen hast ein **tier** punkte 2 wir haben 8.2 millionen tiere in der sch Katze das sagt eine Hause auf 2 habe minimum ein **Tier**. Die Numer 2 ist ein Podium mit ein Hund im letz ustiere im Schweiz.In Schweiz 1/2 Hause hat eine **Tier**.1. Sech personen habt 1 Tiere. 2. 8.2 millionen t chweiz besitzt jeder zweite Haushalt mindestens 1 **tier**. 2) In der Schweiz sind 8,2 Millions Tieren. 3) 3 e Erste Punkt sagt dass 1/2 Hause hat minimum ein **Tier**. Die Zweite sagt dass das Tier das mehr Personen se hat minimum ein Tier. Die Zweite sagt dass das **Tier** das mehr Personen hat ist Fische. Die Dritte sagt nicht sparen.In der Schweiz gibt es 8.2 m. Hause **Tier**. Eine Hause / zwei Hause hat minus ein Tier. Zum Hause Tier. Eine Hause / zwei Hause hat minus ein **Tier**. Zum beispiel, wenn wir 10 personen wären, hätte enn wir 10 personen wären, hätte 5 personen ein **Tier** zu hause. Einen hund ist billiger einen katze, abe



Figure 4 Excerpts from SWIKOweb based on SWI02 (graph about pets) in DaF texts: concordances (top), word cloud (bottom left), and frequency list (bottom right).

Application and transfer

Finally, in order to apply their knowledge, students can be prompted to write short texts. Out of the eight tasks in SWIKO, learners used negations most often when creating a self-portrait or discussing a list of the most important inventions, though other tasks might provide even more opportunities to use the target construction. Again, several scaffolding options are conceivable: The teacher can offer an exemplary text, or students can also be encouraged to use the mind map created earlier (Übungsblatt Negation 2). Furthermore, when creating the worksheets for the introduction and practice (Übungsblatt Negation 1-4), the teacher could select concordance lines only based on the same type of task, and these can then serve as additional templates.

Conclusion

Despite a large increase in corpus linguistic research studies, corpora have yet to find their way into the foreign language classroom. We aimed to bridge this gap by discussing two scenarios on how our corpus-linguistic research findings based on the rich and authentic Swiss Learner Corpus SWIKO can be used in foreign language education.

First, in teacher training, the productions can serve as an illustration of learners' abilities at the end of mandatory schooling, particularly in combination with our findings on task-based differences regarding the length, complexity, and accuracy of learner language. Second, our analysis of frequent errors in foreign language productions shed light on particularly challenging structures, while the language of schooling sub-corpus can serve as a peer-reference in the development of corresponding material. We exemplified this process focusing on negation in German, offering differentiated teaching material suitable for the secondary school classroom.

We hope that our contribution encourages readers take a leap and consider using learner corpora such as SWIKO in their classroom – whether as an authentic resource to illustrate task-based differences in learner productions or to introduce and consolidate a lexical or grammatical phenomenon through an autonomous and collaborative discovery approach.

References

- Alexopoulou, T., Michel, M., Murakami, A., & Meurers, D.** (2017). Task Effects on Linguistic Complexity and Accuracy: A Large-Scale Learner Corpus Analysis Employing Natural Language Processing Techniques. *Language Learning*, 67(51), 180–208.
- Bertschy, I., Cuenat, M. E., & Stotz, D.** (2015). *Lehrplan Französisch und Englisch*. Passepartout - Fremdsprachen an der Volksschule.
- Boulton, A., & Cobb, T.** (2017). Corpus Use in Language Learning: A Meta-Analysis. *Language Learning*, 67(2), 348–393.
- Council of Europe.** (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.
- Council of Europe.** (2020). *Common European framework of reference for languages: Learning, teaching, assessment: companion volume*. Council of Europe Publishing.
- Crosthwaite, P. (Ed.).** (2020). *Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners*. Routledge.
- Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C.** (2020). *Task-based language teaching: Theory and practice*. Cambridge University Press.
- Flinz, Carolina** (2021). Korpora in DaF und DaZ: Theorie und Praxis. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 26(1), 1–43.
- Gilquin, G., & Granger, S.** (2010). How can data-driven learning be used in language teaching? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge Handbook of Corpus Linguistics* (pp. 359–370). Routledge.
- Granger, S.** (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, 1(1), 7–24.
- Granger, S. and C. Tribble** (1998). Learner corpus data in the foreign language classroom: form-focused instruction and data-driven learning. In S. Granger (Ed.), *Learner English on Computer* (pp. 199–209). Routledge.
- Housen, A., Kuiken, F., & Vedder, I. (Eds.).** (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*. John Benjamins.
- Johns, T.** (1997). Contexts: The background, development and trialling of a concordance-based CALL program. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 100–115). Longman.
- Karges, K., Studer, T., & Hicks, N. S.** (2022). Lernersprache, Aufgabe und Modalität: Beobachtungen zu Texten aus dem Schweizer Lernerkorpus SWIKO. *Zeitschrift für germanistische Linguistik*, 50(1), 104–130.
- Karges, K., Studer, T., & Wiedenkeller, E.** (2019). On the way to a new multilingual learner corpus of foreign language learning in school: Observations about task variation. In A. Abel, A. Glaznieks, V. Lyding, & L. Nicolas (Eds.), *Widening the Scope of Learner Corpus Research. Selected papers from the fourth Learner Corpus Research Conference* (pp. 137–165). Presses universitaires de Louvain.
- Lemnitzer, L., & Zinsmeister, H.** (2015). *Korpuslinguistik: Eine Einführung*. Narr Francke Attempto.
- Lüdeling, A., & Hirschmann, H.** (2015). Error annotation systems. In F. Meunier, G. Gilquin, & S. Granger (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 135–158). Cambridge University Press.
- McEnery, T., & Xiao, R.** (2011). What Corpora Can Offer in Language Teaching and Learning. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning* (pp. 364–380). Routledge.
- Meunier, F.** (2019). Data-Driven Learning: From Classroom Scaffolding to Sustainable Practices. *ELLE*, 2, 423–434.
- Meunier, F.** (2020). Introduction to learner corpus research. In N. Tracy-Ventura & M. Paquot (Eds.), *The Routledge Handbook of Second Language Acquisition and Corpora* (pp. 23–36). Routledge.
- Römer, U.** (2008). Corpora and language teaching. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics. An International Handbook* (pp. 112–130). Mouton de Gruyter.
- Römer, U.** (2009). Corpus research and practice: What help do teachers need and what can we offer? In K. Aijmer (Ed.), *Corpora and Language Teaching* (pp. 83–98). John Benjamins.
- Schmid, H.** (2013). *TreeTagger—A Language Independent Part-of-speech Tagger* (3.2) [Computer software]. <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger>
- Studer, T.** (2023). Handlungsorientierung im Unterricht? Ja, klar. Aber was und wie genau eigentlich, warum und was kommt dabei heraus? In S. Demmig, S. Reitbrecht, B. Sorger, & H. Schweiger (Eds.), *IDT 2022: Beiträge zur Methodik und Didaktik von Deutsch als Fremd*Zweitsprache*. (pp. 19–29). Schmidt.
- Studer, T., & Hicks, N. S.** (2022). *The interplay of task variables, linguistic measures, and human ratings: Insights from the multilingual learner corpus SWIKO*. European Second Language Acquisition Conference, Fribourg.
- Tribble, C.** (2008). From Corpus to Classroom: Language Use and Language Teaching. *ELT Journal*, 62(2), 213–216.
- Vyatkina, N.** (2020). *Incorporating Corpora: Using corpora to teach German to English-speaking learners* [Online instructional materials]. University of Kansas Open Language Resource Center. <https://corpora.ku.edu/>
- Weiss, Z., Hicks, N. S., Meurers, D., & Studer, T.** (2022). *Using linguistic complexity to probe into genre differences? Insights from the multilingual SWIKO learner corpus*. Learner Corpus Research Conference, Padua.
- Wicher, O.** (2019). Data-driven Learning in the Secondary Classroom: A Critical Evaluation from the Perspective of Foreign Language Didactics. In P. Crosthwaite (Ed.), *Data-driven Learning for the Next Generation: Corpora and DDL for Pre-tertiary Learners* (pp. 31–46). Routledge.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y.** (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. University of Hawaii Press.

Übungsblätter

Download «Übungsblätter»:
<https://ifm-swiko.unifr.ch/publications>

Übungsblatt Negation 1a: Negation im Deutschen

Im Deutschen gibt es v.a. zwei Möglichkeiten, Sätze zu verneinen (Option 1 und 2). Wann wird welche Option gebraucht? Schau dir die Sätze an und versuche, eine Regel daraus abzuleiten.
 Tipp: Achte besonders auf das **hervorgehobene Wort in der Mitte** sowie das **erste Wort nach dem hervorgehobenen Wort**.

Option 1:

Aber, Ich bin Vegetarier, dass heisst ich esse gar **kein** Fleisch. Meine Schwächen sind, manchmal nicht wie ein Grosseisler. Skiferien finde ich toll. Ich bin **kein** Fan von Jungendherbergen. Ja ich bin einverstanden wenn Spinnen hab ich angst, kleinere Spinnen sind **kein** Problem. Ich finde, dass die Elektrizität an erste eine Mensch besitzt eine Katze der andere hat gar **kein** Tier. Es ist auch möglich das derjenige mit dem T... Ich mag wenn ich die Frage zum Essen beantworte **keine** Pilze. Die Konsistenz und der Geschmack ist nicht spiele ich Fussball oder Computerspiele. Ich habe **keine** Lieblingsmusik. Ich höre verschiedene Musikarten t auf der Liste stehen, denn ohne den, hätten wir **keine** Bücher schreiben können. Das Fotoapar ist auch

Option 2:

Aber was ich genau machen möchte, weiss ich noch **nicht** genau. Ich liebe Dessert, vor allem wenn das Dessert gefährlich sind und ekelhaft. Etwas was ich auch **nicht** gerne habe, ist wenn es in den Bergen sehr stark ane. Mit dem Punkt Ausflüge in die Berge bin ich **nicht** einverstanden, weil ich mega gerne in die Berge f n. weil in den Bergen ist es immer sehr schön und **nicht** so viele Leute wie in Städten. Mit dem Punkt Städ kt städtereisen bin auch einverstanden, aber auch **nicht**, weil eine städtereise zu machen ist auf einer Se ndert haben. Denn die Welt ohne Elektrizität wäre **nicht** so cool und man hätte nicht so viele elektrische

Wie wird die Negation gebildet? Schreib die Regel und ein Beispiel dazu auf.

Regel 1: kein/e + _____ Beispiel: _____

Regel 2: nicht + _____ Beispiel: _____

Babylonia SWIKO - Corpora in foreign language education 12

Übungsblatt Negation 1b

Im Deutschen gibt es v.a. zwei Möglichkeiten, Sätze zu verneinen (Option 1 und 2). Wann wird welche Option gebraucht?

- Sortiere die Streifen in zwei Gruppen und klebe sie auf.
- Wie wird die Negation gebildet? Schreibe die Regeln und ein Beispiel dazu auf.

Regel 1: _____ + _____ Beispiel: _____

Regel 2: _____ + _____ Beispiel: _____

Babylonia SWIKO - Corpora in foreign language education 13

Material für Übungsblatt 1b ausdrucken und in Streifen schneiden

Aber, Ich bin Vegetarier, dass heisst ich esse gar **kein** Fleisch. Meine Schwächen sind, manchmal nicht wie ein Grosseisler. Skiferien finde ich toll. Ich bin **kein** Fan von Jungendherbergen. Ja ich bin einverstanden wenn Spinnen hab ich angst, kleinere Spinnen sind **kein** Problem. Ich finde, dass die Elektrizität an erste eine Mensch besitzt eine Katze der andere hat gar **kein** Tier. Es ist auch möglich das derjenige mit dem T... Ich mag wenn ich die Frage zum Essen beantworte **keine** Pilze. Die Konsistenz und der Geschmack ist nicht spiele ich Fussball oder Computerspiele. Ich habe **keine** Lieblingsmusik. Ich höre verschiedene Musikarten t auf der Liste stehen, denn ohne den, hätten wir **keine** Bücher schreiben können. Das Fotoapar ist auch

Aber was ich genau machen möchte, weiss ich noch **nicht** genau. Ich liebe Dessert, vor allem wenn das Dessert gefährlich sind und ekelhaft. Etwas was ich auch **nicht** gerne habe, ist wenn es in den Bergen sehr stark ane. Mit dem Punkt Ausflüge in die Berge bin ich **nicht** einverstanden, weil ich mega gerne in die Berge f n. weil in den Bergen ist es immer sehr schön und **nicht** so viele Leute wie in Städten. Mit dem Punkt Städ kt städtereisen bin auch einverstanden, aber auch **nicht**, weil eine städtereise zu machen ist auf einer Se ndert haben. Denn die Welt ohne Elektrizität wäre **nicht** so cool und man hätte nicht so viele elektrische

Babylonia SWIKO - Corpora in foreign language education 14

Übungsblatt Negation 2: Kollokationen

Negationen kommen oft in typischen Wort-Verbindungen, so genannten Kollokationen vor.
 Sammelt typische Kollokationen zu den zwei Begriffen **nicht** und **kein**.

nicht (oo) gut

keine (guter/schlechte) Idee

Babylonia SWIKO - Corpora in foreign language education 15

Übungsblatt Negation 3: nicht oder kein/e?

1) Ergänze das fehlende Wort in der Lücke.

tan höre ich Klassik und Soundtracks. Ich bin mir _____ sicher, ob es wirklich Angst ist, aber es läuft m
 ein. Das macht so viel Spaß. Ich finde die Liste _____ unheimlich lustig. Das macht mich mit Partizip
 wird man auch sagen könnte, dass es über Computer _____ Interesse gibt. Außerdem ist die Belle viel zu n
 ist eine wichtige Erfindung, ohne das hat würde es _____ Wissen oder Autos und weitere Erfindungen geben. I
 ste aus meiner Meinung nach viel weniger oder gar _____ Fleisch habe Essen, dass es gibt auch lockere G
 e werden immer mehr. Partizip sind sicher auch _____ möglich. Ferien in der Schweiz finde ich auch ge
 e meistens sind doch doppelt so viel Spaß. Was ich _____ gerne mache, ist auf den Bauernhof die Ferien zu
 Wanderferien finde ich gar schrecklich. Ich bin _____ damit einverstanden, dass Österreichern langweilig
 ist ist mein Lieblingsort Fremden. Ich mag _____ Menschen, da sie gefährlich sind und ehrlich. E
 en kleinen finde ich eher unangenehm, da sie oft _____ mehr machen sind. Mit dem Partizip mit Freude
 Glückseligkeit, was hat durch die Sprachbarriere _____ Freude. Die Sprache wird zu einem Hindernis. Man
 am 2 Stunden Hausaufgaben machen, haben sie dann _____ Freizeit mehr. Es wäre aber gut, später in der Sc
 reitag um Mittag ist gut. Aber so lange will ja _____ darüber in der Schule bleiben. Dann habe ich lieb

2) Bilde die Negation mit den vorgegebenen Wörtern und halte es in den Kästen fest.

die Idee - gut - sicher - die Zeit - mehr - die Lust - das Problem - genau
 gerne - das Lieblingsessen - einverstanden - so toll - das Haustier

kein/e
keine (gute oder schlechte) Idee

nicht
nicht gut

Übungsblatt Negation 4: richtig oder falsch?

1) Wie wird die Negation in deiner Schulsprache gebildet? Und in anderen Sprachen, die du kennst?

Diskutiert eure Beobachtungen in der Klasse. Vergleicht sie dann mit der Negation im Deutschen.

2) Unten siehst du Textauschnitte von Sekundarschüler/innen, die Deutsch als Fremdsprache lernen. Korrigiere die Negationsfehler - aber Achtung! Nicht alle Textauschnitte enthalten Negationsfehler. Entscheide zuerst, ob es richtig oder falsch ist und korrigiere nur falls nötig.

Beispiele:

Die Ferien in der Schweiz sind nicht gut weil es keine interessante Sachen hat. Die Ski ferien sind cool
 Schule. Und ich mag eat. Ich weiss es nicht. Ich plane Lust zu ending die cover in 17 h 20. Ich warte ni

Meine Handy, weil mein Samsung ist kaputt. Ich habe nicht ein Lieblingsbuch aber ich lese viele Buch und es

rgi. Ich mague Volleyball am besten. Ich weise es nicht. Mein Lieblings aktivität ist in die shop gegangen

vi lesen oder mit Computer gut arbeiten. Ich habe kein Haustier aber am Sommer bekomme ich ein Hund. Ich

ant. 17h30 ist nicht gut für aktivitäten ich habe nicht Zeit. Die pause ist kurz. Ich habe nicht für wei

er Das ist mehr important als die Brille. Es gibt nicht Autos und es ist sehr important in die Planet. Ich

meis, weil wir läte aus ur Hause sind. Wir haben nicht Freunden für spielen, diskutieren ect. Wir sind i

Bett weil man Hausaufgaben macht. Wir haben auch keine Lust für Aktivitäten mit freunden oder familie. D

, weil ist mit freunds. Ich denke dass camping ist nicht cool, weil es is altmodisch. Ich mag nicht pferdef

r ist super wir müssen hören Musik aber wir haben nicht elektricité wir müssen nicht haben Computer nicht

se zeit für essen Mittag. Die pose für mittag ist nicht eine gute Idee. Wir können schlafen aber das ist