Universität Tübingen

Seminar für Sprachwissenschaft

Wilhelmstraße 19

72074 Tübingen

M.A. Thesis in Computational Linguistics

Android App Development for Language Learning: Combining input and interaction in learning German through song lyrics

Alina Baranova

alina.baranova@student.uni-tuebingen.de

First Examiner and Supervisor:	Prof. Dr. Detmar Meurers
Second Examiner:	Johannes Dellert, PhD

Hiermit versichere ich, dass ich die vorgelegte Arbeit selbstständig und nur mit den angegebenen Quellen und Hilfsmitteln einschließlich des WWW und anderer elektronischer Quellen angefertigt habe. Alle Stellen der Arbeit, die ich anderen Werken dem Wortlaut oder dem Sinne nach entnommen habe, sind kenntlich gemacht.

_____

(Alina Baranova)

Abstract

The thesis describes the development of an Android app for learning German grammar through listening to German music. Song texts and song audios are used as input in the language of learning necessary for language acquisition. The grammatical constructions that the learners of German experience the most difficulties with are detected and extracted from song texts. These grammatical constructions are used for input enhancement — they are highlighted in song texts to draw the attention of the learner to the structure of constructions. They are also used as basis for the multiple-choice questions implemented in the song texts, which provide the learner an opportunity to practice their grammar skills. A wide variety of automatically generated distractors facilitate the praxis mode of the app.

Zusammenfassung

In dieser Masterarbeit stelle ich die Entwicklung einer App zum Erlernen der deutschen Sprache dar. Als Input für den Zweitsprachenerwerb verwende ich Liedtexte und die zugehörigen Audiodateien. Ich identifiziere diejenigen grammatikalischen Aspekte der deutschen Sprache, die die meisten Herausforderungen für Deutschlernende darstellen, und finde dazu passende grammatikalische Konstruktionen in den Liedtexten. Diese grammatikalischen Konstruktionen verwende ich zum Input Enhancement. In den Liedtexten werden sie hervorgehoben, um die Aufmerksamkeit des Lernenden auf ihre Struktur zu lenken. Die grammatikalischen Konstruktionen sind auch die Basis für praktische Übungen, die die App ermöglicht. Durch Multiple-Choice-Fragen, die in den Liedtexten implementiert sind, kann der Lernende seine grammatikalischen Fähigkeiten verbessern. Eine Vielfalt automatisch generierter Distraktoren ermöglicht die effektive Sprachübung.

Contents

1 Introduction

Second language (L2) learning requires a variety of different components. One of its most essential components is input in the language that is being learnt — the term applies both to the texts written in L2 and to spoken language. However, acquiring the language cannot be ensured without two additional factors — the learner paying attention to certain language aspects of the input and practicing the skills connected to these aspects.

It is important to detect the aspects of language that pose particular challenges to L2 learners. As this thesis is concerned with German as a foreign language, the book "Mitsprache fördern" can be useful (Koller et al., 2013). It introduces the aspects of German that non-native German speakers studying at German schools experience most difficulties with. Another helpful resource is learner corpora. The MERLIN corpus, for example, provides a collection of texts written by learners of a few languages, including German.

To connect the typical challenges for learners of German language and the components necessary for language learning, an app for Android-based mobile devices is developed. The app includes lyrics of German songs that are synchronized with their audio, and provides the student with a highlight mode and a game mode. Highlighting certain language aspects is a kind of input enhancement that helps the learner to acquire declarative knowledge about the highlighted constructions. The game mode involves practical exercises that facilitate transforming the declarative knowledge into the procedural knowledge. The exercises are supposed to be accomplished while listening to songs, which increases the user's involvement in the process of L2 learning. That helps to maintain the student's motivation in the process of learning. Including the songs popular among learners of German and providing the user with an opportunity to choose songs of their taste also add to the user's motivation.

The thesis is structured as follows. The next section provides the theoretical background for the current work, describes the existing systems facilitating L2 learning through lyrics and lists the limitations of the systems. Section 3 covers the

details of our approach and discusses the tasks that have to be accomplished in order to successfully develop the app for language learning. Section 4 describes the process of compiling a corpus of song texts and the synchronization of texts with their audio. The grammatical topics implemented in the app and the NLP analysis carried out in order to extract the corresponding grammatical constructions from the lyrics are discussed in section 5. Section 6 is concerned with different strategies of generating distractors for the multiple-choice questions used in the praxis mode of the app. The resulting application and its functioning are described in section 7. The last section introduces various strategies of improving the app that can be exploited in the future work.

2 Related Work

This section is devoted to discussing the theoretical background behind the second language acquisition. It also describes and compares the existing systems that offer the user a possibility to learn languages through song lyrics.

2.1 Theoretical background

Successful learning of a L2 requires many components — among others, the linguistic data produced by native speakers of the language is necessary. Krashen (1985) suggested that such *comprehensible input* plays the central role in the second language acquisition. As he formulated in the Input Hypothesis, if the input provided to the student is a bit more complex than his/her current level of competence, perceiving the input will automatically lead to learning the grammar. The hypothesis was based on drawing an analogy between acquiring the first and the second language.

The later research, however, has pointed out that even though the input is important, it is not sufficient for learning a language. Even for young L2 learners, acquisition of the L2 through oral messages with interlocutors and written texts surrounding the learners led only to minimal grammatical progress (Ortega, 2013). Therefore, in case of an app for language learning, providing the students only with German songs and their texts would not result in a significant improvement of students' language performance.

Schmidt (1995) suggested that the acquisition of the second language takes place when the learner pays attention to the input – namely, notices different aspects of phonology, morphology, pragmatics, discourse, etc. Based on the other experiments, Schmidt concludes that perceiving new mental concepts without attending to them may result in their detection, but is not sufficient for their learning. Ortega (2013) explains the necessity of attention by the fact that it leads to a higher activation level of input in working memory. Through rehearsal, the input stays in working memory even longer and becomes available for subsequent processing and transferring to long-term memory. Thus, the language level can only be improved through listening

to music and reading song texts when the students pay attention to certain aspects of the language while doing it.

Sharwood-Smith introduced a concept of *consciousness-raising* — a way of drawing the learner's attention to some aspect of the target language (Sharwood Smith, 1981). Some years later, however, he stated that using the term *input enhancement* for the concept was more appropriate (Sharwood Smith, 1993). While consciousness-raising assumes that the learner's mental state has been altered by input, input enhancement simply states the fact that the input has been changed so that the leaner pays attention to a certain linguistic aspect. In the latter case, it is not clear if the input has affected the learner and whether he/she has driven any conclusions from it. Smith distinguishes different kinds of input enhancement. Firstly, one could enhance positive or negative input. Positive enhancement implies making correct forms salient — for instance, making all endings of English verb forms of 3rd person, singular number and present tense bold. Negative enhancement, on the other hand, means signalling the learner which forms he/she used incorrectly. An L2 text written by the learner in a text editor and corrected by a built-in spellchecker could serve as an example of negative enhancement. Secondly, input enhancement can differ in terms of elaboration degree. Using color-coding in text or intonation in speech would be assigned a low degree of elaboration, while using metalinguistic terminology for explaining a construction would involve a higher degree.

Being aware of some aspects of language, however, constitutes only declarative knowledge about these aspects — that is, *knowing **that***. With practice, the declarative knowledge gets converted into the procedural knowledge — *knowing **how***. Practice helps the conversion by making the links in long-term memory corresponding to the new knowledge stronger (Ortega, 2013).

Subsequent practice may lead to making the procedural knowledge automatized. Automatized knowledge can be characterized by low reaction time and low error rate. One should note that practice is skill-specific. If we compare, therefore, production and comprehension skills in a second language, practicing production skills may lead to acquiring some declarative knowledge about comprehension, but the procedural knowledge would not be transferred from one type of skills to another (Dekeyser, 2007). Exercises connected to song texts are, however, connected both to comprehension and production skills. Practicing comprehension skills involves listening to songs and adopting to aritists' pronunciation, as well as perceiving written song texts. Completing the texts, on the other hand, could be considered a way of practicing the production skills.

Another important component of the successful L2 learning is motivation. While most of the research on motivation in L2 between the 1960s and 1990s was focused on the learners' perception of the L2, the L2 community and the L2 culture, the researchers in the 1990s agreed that the social dimension alone does not explain many other motivational factors studied in the field of educational psychology. They, therefore, wanted to *close the gap between motivational theories in educational psychology and in the L2 field* (Dörnyei, 2001). Dörnyei held similar views. Building on his framework presented in 1994, he introduced a process model of learning motivation in the L2 classroom in his book "Motivational Strategies in the Language Classroom" (Dörnyei, 2001). The model assumes that the learner experiences different types of motivation before, in the process of and after the action — thus, in the preactional, actional and postactional stage of acquiring the L2. The preactional stage includes such motivational influences as the expectancy of success, the values associated with the learning process and its outcomes and the learner's attitudes towards the L2 and its speakers. It is important, however, not only to generate motivation in the initial stage, but also to maintain it during the learning process. One of the strategies the author suggests involves making tasks for practicing language aspects more enjoyable. That could be accomplished in a number of ways, including introducing content interesting for the learner and personalizing the exercises. Exploiting music and song texts for language learning meets both of these requirements. Thus, if the songs are collected according to the music taste of the majority of L2 learners, the song texts would be entertaining for a great number of students. Apart from this, the freedom of the learner to choose his/her favorite songs would make him/her more motivated to accomplish the exercises connected to the songs.

To sum up, song texts in L2 can serve as the input necessary for language acquisition. To improve the learner's language level, however, these texts should be enhanced in a way that the learner pays attention to a number of language aspects. By turning the texts into exercises involving these language aspects, the learner would be able to make his/her comprehension and production skills better. During the exercises, the entertaining content would help to keep the learner's motivation high.

2.2 Existing systems

Only websites and applications that involve exercises based on songs texts were examined. Thus, collections of song lyrics and their translations were not taken into

account.

The following systems were analyzed: *LyricsTraining*[1], *FluentU*[2], *LyricsGaps*[3], and *Language Zen*[4]. All the four systems have a web version and a mobile version. Although the creators usually invest in one of the versions more, features of both web and mobile versions are discussed.

Different aspects of systems were analyzed: possible modes of discovering lyrics, type and structure of exercises for memorizing new words, language levels that the system offers, translation of separate words and whole lyrics into English. Moreover, the number of languages supported by systems and features not connected to song lyrics were mentioned. The systems that have more popular Android apps were highlighted first.

2.2.1 LyricsTraining

According to the Google Play[5], an official app store for the Android operation system, the mobile version of *LyricsTraining* is the most popular among the four systems mentioned here, and has more than 500,000 downloads[6]. The system is solely developed for using music lyrics in language learning. When clicking on each song, the user can choose one of the two modes: the karaoke mode or the game mode. In the game mode, a level — beginner, intermediate, advanced or expert — should be selected. The more difficult the level is, the greater the number of blanks in the text is. In the web version, every song lacks either 10%, 25%, 50% or 100% of the words. In the mobile version, the number of words transformed into questions differs from song to song, but in the expert level the user has to fill in all words in the lyrics, as in the web version. Such task is very unlikely to be accomplished, regardless of the language level of the user. Moreover, it hinders one of the main goals of learning vocabulary through song lyrics — learning new words in context. Therefore, for the expert level, it would be reasonable to decrease the number of words that are transformed into questions.

In all the other three levels in the mobile version, the number of words extracted from the text seems to be dependent on the pace of the song. It is more comfortable than filling in 10%, 25% or 50% of song lyrics, as in some very fast songs it might

---

[1] https://lyricstraining.com
[2] https://www.fluentu.com
[3] https://www.lyricsgaps.com
[4] https://www.languagezen.com/en
[5] https://play.google.com/store
[6] Accessed: June 28, 2019

be too hard for the user to accomplish a higher level, which can possibly lead to his/her frustration.

The type of exercises in the mobile and in the web version are different: it is multiple choice questions and fill-in-the-blank questions accordingly. For every multiple choice question in the mobile version of *LyricsTraining*, three distractors are available, all of them the words from the other blanks in the song. On the one hand, the user can easily guess the word while hearing it, as the distractors are not especially chosen to sound similarly to the right answer. On the other hand, the user is less stressed, as he/she needs less time to make a decision about the next word, because three out of four options from the previous question stay the same in the next one. Apart from that, it is programmatically simpler to use words from other blanks as distractors, as there is no need to obtain them from an external source. Moreover, due to the fact that the distractors do not have to be picked beforehand, random generation of blank spaces is possible.

Fill-in-the-blank questions in the web version are considerably harder than multiple choice questions, especially when a singer uses slang, and a song is rather fast. Compared to the mobile version, questions in the web version appear to be less sensible and thought-out and might more easily lead to the user's dissatisfaction with his/her result.

In the game mode, both in the web and in the mobile version, a music video is synchronized with the lyrics, and a certain time limit is implemented — the user is able to choose the option or type in the answer only as long as the line with a blank is sung. If the user did not answer correctly the first time, there is a possibility to hear the line again or to see the right answer and continue listening to the song.

Apart from the game mode, the user can choose a karaoke mode. In this mode, as in the game mode, a music video is synchronized with the lyrics, so that the line sung at the moment is highlighted. Unlike in the game mode, however, user does not have to fill in any words and can switch the song backward and forward. This mode is suitable for either hearing the song for the first time, before the game mode, or afterwards, so that the user can memorize the new words he/she has encountered.

Unlike some other systems, *LyricsTraining* does not allow clicking on words in a song to obtain their translations and does not offer translations of the whole lyrics. However, there is a section *My vocabulary* in the mobile version, where separate words can be translated into 58 other languages through the Microsoft Translator.

Even though only the German version of the *LyricsTraining* was tested, the system

is offered in twelve other languages, among them e.g. Japanese and Turkish.

2.2.2 FluentU

The system with the second popular Android app is *FluentU*, which has more than 100,000 downloads[7], according to the Google Play. Music videos is only one of the resources for language learning the system offers. For the first two levels, Beginner 1 and Beginner 2, audios on grammar topics are available. For all the levels, videos on different topics (e.g. *Arts and Entertainment* and *Science and Tech*) are accessible. The videos have different formats, too, such as *Mini-movies* or *Talks/Speeches*. *Music videos* is one of them.

There are three possible ways one can work with a music video. First, there is a dialogue mode, where the user can read through song lyrics and its line-by-line translation into English. The vocabulary mode contains the words that are taken from the song text and are supposed to be learnt by the user through a quiz. The last mode is playing the video. In this mode, lyrics and videos are synchronized, and the user can go backward and forward. As *FluentU* is one of the systems that allow translating separate words and phrases in the song text, the user can click on any word in the line that is being sung at the moment and acquire the word's translation, while the video is automatically paused. The translations are very detailed, they include a part of speech, some specific characteristics for this part of speech, such as gender for nouns, an audio with a native speaker pronouncing the word and general example sentences. In the web version, the translations are supplemented with example sentences from the song and a suitable picture.

In order to train the vocabulary from the song, the user can do a quiz. There are three types of the exercises: multiple choice questions, fill-in-the-blank questions and ordering. Multiple choice questions include, for example, choosing a right translation of a sentence from English into German. An example of a fill-in-the blank questions would be selecting the right word to complete the sentence. Finally, ordering is concerned with putting the words in the right order to get a line from the song. The quiz is done after watching the video, and there is no time limit for answering the questions.

All the music videos are assigned levels. There are six levels in total: Beginner 1 and 2, Intermediate 1 and 2, Advanced 1 and 2. The more difficult the level, the less music videos is available — Advanced 1 and Advanced 2, for instance, have only

---

[7]Accessed: June 28, 2019

two and one music videos accordingly. There are two ways of solving this problem: either more videos can be added or each music video could be adapted for other levels. There are forty videos in total, and there are some learner-specific songs among them as well — for instance, *The Umlaut Song!* or *Numbers 1-10 Singalong*. This kind of songs is probably not that interesting for an average music lover, even though it might be helpful for memorizing the vocabulary.

Apart from German, nine more languages are implemented, including, for example, Korean and Russian.

Compared to *LyricsTraining*, *FluentU* uses a more complicated quiz format, which can be more helpful for acquiring new words. However, *LyricsTraining* offers interactive learning, which *FluentU* lacks. This way of learning might be more entertaining for the user. A very small number of music videos is available in *FluentU*, but there are some other activities for learning implemented in the system, so it cannot be considered a big disadvantage. The big advantage of *FluentU*, on the other hand, is the possibility to translate words and phrases while listening to the song, as well as an access to translations of the whole lyrics, which is not found in *LyricsTraining*.

2.2.3 Language Zen

The third most popular Android app belongs to *Language Zen* and has more than 10,000 downloads[8], according to the Google Play. As *FluentU*, *Language Zen* offers other ways of learning the language: the user can choose the courses mode, which contains phrases and vocabulary split into topics, or start learning the most frequent words.

In contrast to other systems, *Language Zen* offers only two languages. In the web version, it is only Spanish for English speakers, and in the mobile version, there is also a possibility to learn English for Portuguese speakers. The creators of the system are planning to implement more languages in the future. Only the Spanish course for English speakers was analyzed.

When the user chooses a song, he/she can either select the play mode or the learn mode. In the play mode, lyrics and audios are synchronized, and the user can play the song forward and backward, as in the other systems discussed so far. *Language Zen* also offers translations of individual words when clicking on them. Apart from that, every line is followed by its translation. Translations of some individual words

---

[8]Accessed: June 28, 2019

(e.g. *se*) are complemented with short grammar notes that explain their usage (e.g. *Passive "se"*). In the web version, word forms are also morphologically analyzed — for example, a verb is assigned number, person and tense in present tense.

The learn mode offers a quiz with multiple choice and fill-in-the-blank questions. The user is asked to translate words and phrases from Spanish to English and vice versa. An interesting feature of the system is the way things that the user learnt or attempted to learn before are included in the quiz even if a song which is new for the user is chosen. There is no time limit for answering questions.

There are many levels of Spanish the user can choose from: beginner, beginner plus, intermediate, intermediate plus, advanced, advanced plus and fluent. There is also a near native level, which is only present in the web version. However, to take any level higher than beginner, the user should pass an assessment test. Beginner level offers only fourteen songs; other levels were not examined.

2.2.4 LyricsGaps

The system that has the least popular Android app among the four systems examined in this section is *LyricsGaps*. According to the Google Play, there are more than 500 downloads[9] of this app. It is important to point out, however, that the web version of the system is much more developed than its mobile version.

As *LyricsTraining*, *LyricsGaps* is fully devoted to learning languages through song lyrics. The exercise mode is shared by both the mobile and the web version. The web version, however, also offers a karaoke mode and a word-to-word mode. The karaoke mode displays the whole text, but, unlike in other systems, the line sung at the moment is not highlighted. The word-to-word mode is rather strange: it is supposed that the user has to click on every word of the song while playing the video.

In the web version, the exercise mode includes texts with multiple choice and fill-in-the-blank gaps. The type of questions depends on the level chosen by the user. Multiple choice questions offer too many distractors — usually, all the words from other blank spaces are used for that. Using only a few distractors, as in *LyricsTraining*, would be better for the user, as it would not confuse him/her that much. In the mobile version, only three distractors are used. However, every line has a blank space that should be filled with one of the four possible options, which results in more gaps than in the web version.

---

[9]Accessed: June 28, 2019

In some sense, one can speak about a time limit in the mobile version, as the user has to type in an answer fast enough to follow the song. If he/she does not type the answer in time, the song does not stop, as, for instance, in *LyricsTraining*. This kind of design might have a negative impact on the results of the user and their satisfaction with the system. In the web version, there is no time limit, and the user can listen to the song as many times as he/she would like.

As has been mentioned before, in the web version the user has the opportunity to choose a level of the exercise: beginner, intermediate or expert. Beginner level exercises include multiple choice questions, and the other two levels offer questions of fill-in-the-blank type, which are objectively harder for the user. No correlation between the level and the number of gaps in the text is found, although there are on average more gaps in songs with longer texts[10].

Even though the system does not provide translations of whole lyrics, the user can acquire translations of separate words by clicking on them. If the word form is different from its lemma, the translation of the word form will be shown — for example, *gab* would be translated into English as *gave*. Apart from English, translations into nineteen other languages are possible. Unfortunately, translation of words is only available in the web version of the system.

LyricsGaps is accustomed to twenty languages, apart from German. Compared to the three other systems, it is the greatest number of languages offered. Among popular languages, such as French and Chinese, more exotic ones are offered — for instance, Indonesian or Catalan.

To sum up, there are two systems that offer exercises while listening to a song, and two systems where exercises should be done after listening. A system can either contain different songs for different levels or adapt each song to several levels. In most of the systems, songs are synchronized with their lyrics, and the user is able to click backward and forward buttons to control the song. The maximum number of languages offered is twenty-one, and the least is two. In three out of four systems, translations of separate words are possible. Two of these systems also contain translations of whole lyrics.

2.2.5 The limitations of the existing systems

It is worth mentioning the limitations that the apps described in this section show. Firstly, exercises connected to song texts are devoted to learning vocabulary only.

---

[10]Five songs from different artists were analyzed.

Even though such approach can be easily scaled up for many languages, the opportunity to exploit an abundance of grammatical constructions encountered in the song texts gets lost. A solution for overcoming this disadvantage would involve selecting a few grammatical constructions that pose challenges to most L2 learners and generating exercises with these constructions.

Secondly, both systems including exercises that the learner should do while listening to a song — *LyricsTraining* and *LyricsGaps* — use other words from the song text as distractors for multiple-choice questions. That might make exercises too easy, especially for learners with a high level of language proficiency. To provide learners with more challenging exercises, words that either sound similar or have a similar meaning to the right option could be extracted from the external sources.

The last observation may not apply to all the apps, as it is not clear how the songs for the apps were selected. One of the opportunities would involve using the data from the music charts. Even though they might seem a good resource of information, they are not based exclusively on the data from L2 learners. The preferences of L2 learners can significantly differ from that of native speakers e.g. the learners can favor slower songs with simpler vocabulary. The best way to acquire information about the kind of songs and artists L2 learners prefer is suggesting them to take a part in a survey, which would consist of filling out a questionnaire about their favorite German artists.

We paid attention to these limitations while developing our app and did our best to overcome them.

3 Our approach

In this chapter, we introduce the tasks that we have to accomplish in order to develop an app for learning German. First, a corpus of song lyrics should be compiled. It is important that the songs are chosen by German language learners, so a survey has to be carried out. Apart from that, songs also should be present on YouTube[1], which we plan to use for playing an audio of songs in the app, and we should have free access to songs' texts.

Since all the existing apps implemented learning vocabulary through song lyrics, we have decided to provide the learners an opportunity to practice their grammar skills with songs. It is, therefore, crucial to choose the grammar topics that the learners will be able to train by using the app. Even though German language poses many challenges for its learners, only the topics that the language learners usually have the most problems with should be selected. The grammatical constructions corresponding to these topics should occur in the song texts and suit to the format of exercises implemented in the app.

After that, computational linguistic means appropriate for extracting the grammatical constructions for the chosen topics should be found. Then, a way to extract the constructions has to be invented and evaluated. We decide to implement a rule-based approach for the extraction of constructions, as the size of the corpus is not sufficient for a machine-learning approach.

We have made the decision to develop exercises that the learner will be able to do while listening to the song, as it increases the learner's involvement in the task. According to Dörnyei (2001), increasing the student's involvement is one of the strategies of maintaining his/her motivation on the actional stage. All the lyrics, therefore, have to be aligned with their audio to make such exercises possible.

The type of exercises best suitable for being accomplished during listening to a song is multiple-choice exercises. Thus, the word that will get replaced with a blank has to be identified for all types of grammatical constructions. Apart from that, different approaches for generating distractors should be developed. For

---

[1] https://www.youtube.com

some grammatical constructions, it is clear what word should be substituted with a blank and what kind of distractors for this word would mislead the user the most. For some, however — in our case, for constructions in the perfect tense and for constructions in the passive voice — the learner corpus has to be used to define the words that lead to most mistakes and to classify the types of mistakes that could be exploited for generating the distractors. Programming language Python[2] is used both for the generation of distractors and the extraction of grammatical constructions.

Last but not least, an app should be designed and implemented. According to the data of June 2019, the Android system is the most popular mobile operating system with the market share of 76.03%. iOS, which is ranked second, reached only 22.04% at that time[3]. Therefore, the Android operating system has been chosen for the app. The app has to include all the songs and make use of the extracted grammatical constructions and the generated distractors, in order to provide the learner's with exercises to improve his/her language level while listening to his/her favorite artists. The app is written with the programming language Java[4] and can be accessed at Github[5].

---

[2]https://www.python.org
[3]https://gs.statcounter.com/os-market-share/mobile/worldwide
[4]https://www.java.com/en/
[5]https://github.com/AlinaBaranova/German_Through_Lyrics

4 Compiling a Corpus of Songs

4.1 Selecting songs

For creating an app for learning German grammar through songs, a set of songs had to be sampled. Consulting German L2 learners allowed to address the issue objectively, independent of our own musical preferences, and receive a number of artists belonging to various genres.

Google Forms[1], one of the platforms for online questionnaires, was used for getting information about favorite German artists of language learners. Students whose native language was not German were asked to name three artists or bands whose songs they would like to see in an application for learning German. They also had to choose an appropriate genre for each of them. There were ten options to choose a genre from: pop, rock, dance/electronic/house, soundtracks, hip-hop/rap/trap, singer/songwriter, classical, R&B, soul/blues, metal. The set of genres was taken from the Music Consumer Insight Report conducted in 2018 by ifpi[2]. To make a list of the ten most popular genres for this report, 19,000 participants from eighteen countries were surveyed. The data received from the questionnaire was based on participants' own definitions of genres, as was the survey for the most popular music genres.

In total, twenty-three people filled out the form, which resulted in a sample of fifty-five unique artists.

Since the app is dependent on YouTube for getting the audio for the songs, the songs were chosen based on this video-sharing platform. It was decided to take a random number of songs — three — performed by every artist. For choosing the songs, the videos by an artist were sorted by the greatest number of view counts. Only songs with official videos of the artist were taken into account. Official lyrics videos — as, for example, a video for *In Deiner Kleinen Welt* by Philipp Dittberner — were also considered. Sometimes, however, it was unclear which videos could

---

[1] https://docs.google.com/forms/
[2] https://www.ifpi.org/news/IFPI-releases-2018-music-consumer-insight-report

be counted official, as an artist or band did not have an official YouTube channel — for instance, that was the case for the band *Ja, Panik*. In such cases, the videos that looked like video-clips were chosen. Another important criterion was choosing songs by an artist or band alone, without other musicians featuring in the videos. Nevertheless, this rule also led to an exception, as most of the songs by Audio88, one of the artists chosen by the German language learners, were made in a collaboration with a rapper Yassin. On this stage of sampling a song set, an artist or a band were excluded if they did not have official videos or at least videos that look like video-clips.

Downloading lyrics from the Internet manually would take a considerable amount of time, that is why ways of doing it automatically have been surveyed. According to the website of Genius[3], it contains *over 25 million songs, albums, artists, and annotations*, while providing an API that can be used for downloading song lyrics. Apart from that, LyricsGenius[4] — a Python client for the API — is available. The Genius.com has been, therefore, chosen as a source of song texts for the app. The following code allows downloading lyrics using song name and artist name:

Listing 4.1: Downloading song lyrics

```python
import lyricsgenius

genius = lyricsgenius.Genius(CLIENT_ACCESS_TOKEN_HERE)
song = genius.search_song(SONG_NAME, ARTIST_NAME)

print(song.lyrics)
```

However, not all songs that have videos on YouTube also have lyrics on Genius.com. To avoid such songs, it was checked that the most popular song on YouTube by an artist is provided with lyrics on the website. If this was not the case, songs of the artist were not included in the sample.

All in all, forty-four out of fifty-five artists acquired with the questionnaire were left, resulting in 131 songs for the sample. Only one song by one of the forty-four artists — *Am ersten Sonntag nach dem Weltuntergang* by Element of Crime — was not featured on Genius. That shows that the strategy to predict which artists have song lyrics on the website by checking the most popular song was rather accurate.

---

[3]https://genius.com/
[4]https://github.com/johnwmillr/LyricsGenius

After acquiring names of songs and their artists, lyrics were automatically downloaded with the code similar to the code shown in Listing 4.1. The only strange result was produced by downloading song *Auf uns* by Andreas Bourani — instead of the song text, one would get a list of songs by different artists. Due to this problem, lyrics for this song were manually copied from the website.

When all the texts were downloaded, some parts of texts which would hinder synchronization of lyrics with audio have been automatically deleted e.g. "[Strophe II]:" or text in brackets in "Ja ja, ich weiß [10x]". Due to examples of the second type, it is very probable that at the stage of synchronizing a song text with its audio some lines will need to be inserted. At this stage, some texts will also have to be slightly modified, as the texts of the official versions of songs are sometimes different from their YouTube versions. The lines that constitute comments of the people that added song texts to Genius.com, usually written after the song text e.g. *-Folgt RapGeniusDeutschland!* will be deleted.

Apart from the changes that were made to the texts while synchronizing them with their audios, only the alterations necessary for extracting grammatical constructions were made. These alterations will be discussed in more detail in chapter 5. Orthographic mistakes in song texts were not corrected.

The lyrics tend to contain many spoken language forms that rarely occur in writing. For example, the form of the definite article of singular number, accusative case and masculine gender could be written as *ein'*, *'n* or *'nen*. As there is no standard orthography for such spoken forms, we accept that there can be a massive variation of their written versions. No form was, therefore, considered the only right form, and none of such colloquial forms were corrected.

Song texts containing swear words were not filtered out, and the swear words were preserved in the texts. We value the ecological validity of language material and hold the opinion that it is important to expose L2 learners to authentic language use.

4.2 Lyrics-to-Audio Alignment

4.2.1 Automatic Alignment

Automatic lyrics-to-audio alignment is a problem that might seem similar to the problem of automatic speech recognition. However, some characteristics of singing voice are very different from the ones of speech, which prevents the usage of speech

recognizer for performing the alignment.

Firstly, acoustic characteristics, such as pitch and loudness, tend to fluctuate much more in singing than in speech. Secondly, sounds in singing, especially vowels, tend to be prolonged or shortened, which results in a larger variation of sound duration compared to speech (Kruspe, 2018). The background music is the third difference between singing and speech. As the spectrum of singing voices gets affected by accompaniment sounds, it makes the process of lyrics-to-audio alignment more difficult (Fujihara and Goto, 2012). For this reason, some studies prefer working with unaccompanied singing rather than with polyphonic music.

Some minor difficulties of recognizing singing voice or aligning it with text also include lyrics-specific vocabulary, which differs from vocabulary used in spoken language (Kruspe, 2018), and incomplete song texts, especially due to omitting interjections such as *oh* or *yeah* (Fujihara and Goto, 2012).

As was briefly mentioned above, there are studies aiming at singing voice recognition e.g. by Kruspe (2018) and Sasou et al. (2005), as well as the ones concentrating on lyrics-to-audio alignment. In this thesis, only works on alignment of singing voices to songs texts are reviewed. There is also some research done on unaccompanied singing (Loscos et al., 1999; Sasou et al., 2005; Gong et al., 2015). However, the task of lyrics-to-audio alignment in polyphonic music is more suitable for aligning songs for the future Android application with their texts, so studies on unaccompanied singing are left out in this evaluation.

The granularity of lyrics-to-audio synchronization can be different — possible fragments of lyrics include phoneme (Mauch et al., 2012), syllable (Iskandar et al., 2006), word, phrase (Fujihara et al., 2006; Dzhambazov et al., 2016), line (Wang et al., 2008; Wong et al., 2007; Mesaros and Virtanen, 2008), paragraph (Lee and Cremer, 2008) and section (Wang et al., 2008), which, depending on the interpretation, can correspond to the paragraph-level alignment. Among the works that are analyzed in this section, synchronization on line level is chosen for the greatest number of studies. This granularity of synchronization is also the best option for the song alignment employed in the future Android application.

As could be expected, most authors choose songs in English to evaluate their approach. There are, however, singular studies that concentrate on other languages, such as Japanese (Fujihara et al., 2006; Fujihara and Goto, 2008), Cantonese (Wong et al., 2007) and Turkish (Dzhambazov et al., 2016). To the best of our knowledge, no systems have been tested on songs in German.

It is unclear how many songs should be included in the test set to provide reliable

results, but, except for one work where the system was tested on three songs only (Iskandar et al., 2006), studies typically include from ten to twenty songs in their test set.

Metrics used to evaluate how successful a system performs the alignment task often differ from one study to another, which makes it difficult to compare results between various studies, especially if the granularity of synchronization is also different. Evaluation metrics include average displacement error (Wang et al., 2008; Lee and Cremer, 2008; Mesaros and Virtanen, 2008), average word error rate (Iskandar et al., 2006) and average accuracy (Fujihara et al., 2006; Fujihara and Goto, 2008; Mauch et al., 2012; Dzhambazov et al., 2016), with average accuracy being the most popular metric. Among the systems evaluated with the accuracy metric, the best one achieves accuracy of 0.88 on a phoneme level (Mauch et al., 2012). If one chooses average displacement error as an evaluation metric, the smallest error[5] would be 0.58 seconds on a line level (Wang et al., 2008). Both systems are tested on English songs.

Unfortunately, even though the two systems show rather impressive result, they are not publicly available. Moreover, the accuracy of synchronisation is crucial for the app, as it ensures interactive learning, and even an error of half-a-minute would make a difference. Apart from that, if some of the methods used in the systems are language-specific, the systems will not be applicable to songs in German. In addition, the initial sample of songs for the app contains only 131 songs, so the manual alignment would not be too time-consuming. Due to these reasons, manual lyrics-to-audio alignment was preferred over automatic alignment.

4.2.2 Manual Alignment

There are several programs that support the process of manual lyrics-to-audio alignment.

4.2.2.1 GNMIDI

This program is only available for the Microsoft Windows operating system, and offers a possibility to manually align half of the lyrics in its free version[6]. The user pastes the text of the song as a whole and aligns the lines of the lyrics to the audio

---

[5]A mean error of 7.34 and 9.03 seconds has been achieved in (Kruspe, 2018), but it is hard to compare this result to other works due to their different granularity.
[6]Demo version of GNMIDI 3.17 was tested.

by pressing the *Enter* key when a line is sung. In this way, a file of LRC format is created, where every line of the lyrics is assigned the time when it is sung (see Figure 4.1[7]). In files of this format, the time is usually written as **[mm:ss.xx]**, where **mm** stands for minutes, **ss** — for seconds, and **xx** — for hundredths of seconds.

```
[00:12.00]Line 1 lyrics
[00:17.20]Line 2 lyrics
[00:21.10]Line 3 lyrics
...
[mm:ss.xx]last lyrics line
```

Figure 4.1: LRC Format

When the user exports the resulting file from GNMIDI, however, no hundredths of seconds are saved in the file. The karaoke demonstration in the program (see Figure 4.2) also relies on times without hundredths of seconds.

Die Tränen greiser Kinderschar
• Ich zieh sie auf ein weißes Haar
Werf in die Luft die nasse Kette
Und wünsch mir, dass ich eine Mutter hätte
Keine Sonne die mir scheint
Keine Brust hat Milch geweint
In meiner Kehle steckt ein Schlauch
Hab keinen Nabel auf dem Bauch
Mutter, Mutter
Mutter, Mutter
Ich durfte keine Nippel lecken
Und keine Falte zum Verstecken
Niemand gab mir einen Namen

Figure 4.2: GNMIDI Karaoke Demonstration (song *Mutter* by Rammstein)

Apart from that, GNMIDI does not support the UTF-8 format, which is important for lyrics in German.

---

[7]Example is taken from `https://en.wikipedia.org/wiki/LRC_(file_format)`

4.2.2.2 SYLT Editor

SYLT Editor is a free software, and versions both for Microsoft Windows and Linux are available[8]. In contrast to GNMIDI, the program supports UTF-8 encoding. However, loading a LRC-format file with lyrics failed, so the song text had to be entered line by line. After the user assigns times to lines of the lyrics by pressing the *Space* button when a line is sung, the text should be exported to a file of the LRC format. Unfortunately, exporting a file, as well as importing a file, was not possible due to an occurring error.

4.2.2.3 Megalobiz: LRC Maker & Generator Online

LRC Maker & Generator Online is a website[9] where one can align a song text to its audio after loading an MP3-file and pasting in the text of the song. When a line is sung, the user must click on this line for it to be assigned the time. The user can also check if he/she aligned the lines in a right way by playing the song again — the lines of the lyrics will be highlighted according to assigned times. To save the file in LRC format, the user should press the button *Save*, which results in a UTF-8 encoded file on his/her computer.

To sum up, LRC Maker & Generator Online is the best choice for manual lyrics-to-audio aligning, as it saves the hundredths of seconds in time tags, in contrast to GNMIDI, and is able to export the resulting file, unlike SYLT Editor. Moreover, the website is completely free to use. The website was used for tagging song texts used in the Android app.

---

[8]SYLT Editor 1.1 was tested for this section.
[9]https://www.megalobiz.com/lrc/maker

5 Identifying the target constructions

5.1 Which language constructions are relevant?

There is a variety of difficulties German L2 learners encounter while acquiring the language. Some of the typical mistakes and misunderstandings might also be connected to the learners' language background and be an effect of language transfer (Odlin, 1989). The book "Mitsprache fördern" (Koller et al., 2013) provides a number of topics that present challenges for non-native speakers of German studying at German schools. The set of topics takes into account different language backgrounds of students.

Not every topic could easily result into an interactive exercise which the user can accomplish while listening to a song. For this reason, such topics as *Verb positioning in main and subordinate clauses* or *Separable and inseparable verbs* were not included in the app. The following topics were considered best suitable for the format of exercises: *Prepositions and their cases*, *Verb conjugation*, *Verb with prefixes* and *Passive voice*.

5.1.1 Prepositions and their cases

Native speakers usually intuitively choose the right case for words in a prepositional phrase. Learners of German, however, have to memorize the rules on case usage and practice them with suitable exercises (Koller et al., 2013). Another characteristic of German language that makes the choice of the right case even more difficult is the importance of the context for some prepositions (Koller et al., 2013). Such prepositions can be used both with the accusative and the dative cases — for example, *in* and *im* (*in dem*) in 1.

(1) a. Der Igel ist **in den** Garten gegangen.

b. **Im** dunklen Garten faucht und schmatzt es laut.[1]

---

[1]Both examples are taken from "Mitsprache fördern" (Koller et al., 2013)

Even though the exercises offered in the section *Prepositions and their cases* were limited to local and temporal prepositions, no such limitation was implemented for the prepositions used in exercises in the app. Not only local (*in* in 2a) and temporal (*nach* in 2b), but also e.g. modal (*mit* in 2c) and causal (*wegen* in 2d) prepositions were taken into account.

(2) a. Ich spring' in Singapur **in** das kalte Wasser[2]

   b. **Nach** all diesen Jahren immer ohne Plan fahren[3]

   c. Also wieder **mit** dem Fahrrad auf die Autobahn[4]

   d. Bist mitten in der Nacht **wegen** mir aufgewacht[5]

Nevertheless, a prepositional phrase in a song had to comply to certain rules. Firstly, a preposition's object had to be either a noun or a pronoun e.g. *Blick* in 3a or *dir* in 3b.

(3) a. Verlier' mich in deinem **Blick**[6]

   b. Ich träum dabei von **dir**[7]

Secondly, if the preposition's object was a noun, there had to be a modifier of the object that illustrates the case that the preposition required. For instance, such modifier could be an article (*die* in 4a), a possessive pronoun (*ihrem* in 4b) or a preposition itself — to be precise, a contraction of the preposition with the definite article (*zum* in 4c).

(4) a. Atemlos durch **die** Nacht[8]

   b. Hätt' von **ihrem** letzten Geld [Die Karte heut' gekauft][9] [10]

   c. Schenk' dir 'n Song **zum** Geburtstag[11]

Therefore, such prepositional phrases as *mit Schaschliksoße* in 5a were not taken into account. In a similar way, constructions like *Mit Bombenlegern* in 5b were

---

[2]Mark Forster: Au Revoir

[3]Clueso: Achterbahn

[4]Alligatoah: Narben

[5]AnnenMayKantereit: Oft gefragt

[6]Lea: Immer wenn wir uns sehn

[7]AnnenMayKantereit: Barfuß am Klavier

[8]Helene Fischer: Atemlos durch die Nacht

[9]Parts of examples surrounded by brackets show the text taken from lines either previous or next to the line in the example. When a sentence takes more than one line, other lines are added to the example, so that the reader could see the structure of the whole sentence.

[10]Frittenbude: Bilder mit Katze

[11]Lea: Immer wenn wir uns sehn

avoided, even though *-n* in *Bomberlegern* implied that the preposition required the dative case.

(5)   a. [Sie reden von den alten Werten] **Mit Schaschliksoße** in den Bärten

  b. [Ich werde niemals wieder U-Bahn fahren] **Mit Bombenlegern** aus Islam[12]

Thirdly, the modifier of the object that illustrated the case had to have a full form. For this reason, for instance, a prepositional phrase *von unser'n autonomen Fans* in example 6 was not considered proper.

(6)   Weil wir uns nicht distanzier'n von **unser'n** autonomen Fans[13]

Finally, no constructions with proper nouns were included, as learning the gender of proper nouns was considered a secondary priority for German L2 learners. In most cases, the gender of a noun was needed to use the right form of its modifiers, as for choosing *die* in the example 7.

(7)   Atemlos durch **die|den|das** Nacht[14]

5.1.2 Verb conjugation

The section on verb conjugation includes a few exercises on the perfect tense, but it is mainly concentrated on the present and the past tenses (Koller et al., 2013). In the app, among the present (*drehst* in 8a), the past (*ließ* in 8b) and the perfect tenses (*hast abgenommen* in 8c), also the pluperfect tense (*hattest vorgetrunken* in 8d) was taken into account. No preference to any tense was exhibited — the frequency of forms in a certain tense depended only on the song texts. It is clear that this topic only addressed finite verbs and did not include infinitives.

(8)   a. Und du **drehst** dich in deiner kleinen Welt[15]

  b. Ich **ließ** die Sonne nie untergehen[16]

  c. Unsere Fotos **hast** du **abgenommen**[17]

---

[12]Both examples were taken from the song "Professoren" by Adam Angst
[13]Zugezogen Maskulin: Alles brennt
[14]Helene Fischer: Atemlos durch die Nacht
[15]Philipp Dittberner: In deiner kleinen Welt
[16]Adel Tawil: Lieder
[17]Revolverheld: Halt dich an mir fest

     d. [Du warst ein bisschen drauf] Und **hattest vorgetrunken** [Vodka mit deinen Freunden][18]

The distinction between irregular and regular verbs[19] deserves special attention in this topic. Although one could argue that more time should be invested into learning irregular verb forms, being exposed to generalizations and learning regular patterns is at least as important. The study by Schmidt (1983) illustrates that well. The subject of study, a native Japanese speaker who moved to Hawaii learned English for three years mainly by speaking it and acquired a rather high level of fluency. His grammatical competence was, however, rather limited — for instance, he was able to use irregular verb forms in past tense, but was unable to make past forms of regular verbs. Therefore, all verb forms have to be present in the app to expose the user to both irregular and regular patterns.

As for prepositional phrases, there is a set of limitations for verb forms. Firstly, forms of *sein*, *haben* and *werden*, as well as modal auxiliary verbs — *müssen*, *können*, *sollen*, *dürfen*, *wollen*, *mögen* — were not extracted. These verbs are highly frequent, so they would certainly dominate if they are not excluded. We, on the other hand, would like to concentrate on lexically rich verbs, as the user might not have come across some of them yet, since their frequency is lower.

Forms of the perfect and the pluperfect tenses were the only exception to the rule — in this case, the whole form including both an auxiliary and a main verb were extracted, as *bin aufgewacht* in the example 9.

(9)   Und heute **bin** ich **aufgewacht**[20]

Secondly, no verb forms that were exclusively used in spoken German were extracted. For example, we did not extract *hab'*, *zeig*, or *geh'n* from lyrics examples in 10.

(10)   a. Ich **hab'** den YouTuber-Look, wie im Buche gedruckt[21]

       b. Deshalb **zeig** ich was ich kann an dem verdammten Mikrofon[22]

       c. Komm, wir **geh'n**[23]

---

[18]Frittenbude: Bilder mit Katze

[19]According to Durrell (2017), both strong and weak verbs can be divided into regular and irregular verbs. Here, the term *irregular* includes all strong verbs and irregular weak verbs

[20]Casper: Auf und davon

[21]Alligatoah: Du bist schön

[22]Haze: Weisch Weisch

[23]Trettmann: Gottseidank

Forms of the perfect tense with spoken versions of auxiliaries were also avoided, for example *hab' erkannt* in the example 11.

(11) Ich **hab'** 'nen Nazi am Geruch erkannt[24]

Since spoken language usually gets acquired during interaction with German native speakers, our goal was to target standard German language, using the app as an intermediate between spoken and written German. Our pedagogical goal was also in line with technical capabilities of the parser used to extract verb forms[25]. As it always split the word on the apostrophe, that sometimes led to the wrongly determined infinitive of a verb. In case of topic *Verb conjugation*, infinitives were necessary for generating distractors.

Due to homonymy of verbs in the imperative mood and verb forms of 1st person, singular number and present tense used in spoken German, in some cases it was not clear if the verb form had to be extracted, even when the surrounding context was taken into account. For example, *lass* in example 12 could be read as *(ich) lass' alle Schatten fort* or as *lass (du) alle Schatten fort*. In such cases, a word was considered a verb in the imperative mood and was, therefore, extracted.

(12) Dein Wort hält Wort, **lass** alle Schatten fort[26]

Thirdly, contractions of *es* with a verb like *hat's geschrieben* in the example 13 were avoided. It was assumed that such contracted forms would draw the learner's attention from the verb form to the contraction.

(13) Aber wär' auch traurig, mein', er **hat's** ja auch für dich **geschrieben**[27]

Fourthly, a verb form was selected even if its prefix was written separably — for instance, *fängt* in example 14. The decision was based on the fact that the verbs that share the same root, regardless of having or not having a prefix, are usually conjugated in the same way. Thus, the form of 3rd person singular of *fangen* is *fängt*, and of *einfangen* — *fängt ein*, or *einfängt*.

(14) Du, dich **fängt** niemand ein[28]

---

[24]Adam Angst: Professoren
[25]On parsers see the section 5.2
[26]Helene Fischer: Phänomen
[27]Audio88: Gnade
[28]Helene Fischer: Phänomen

The last but not the least, if there was several past participles that share the same auxiliary verb as illustrated in example 15, only the first participle together with the auxiliary was extracted. In the highlight mode of the app, it is important to perceive a past participle and an auxiliary verb as one verb form, which would not be possible if all the participles are extracted.

(15)   Du **hast** mich **angezogen**, **ausgezogen**, **großgezogen**[29]

5.1.3 Verb with prefixes

The section on systematic vocabulary expansion presents such methods of word derivation as adding prefixes or suffixes and compounding (Koller et al., 2013). Many exercises in the section are solely devoted to verb prefixing and changes in meaning that these prefixes lead to. Due to these difference in meanings, verbs with prefixes suited to multiple choice questions well — changing the prefix of a verb easily produced distractors for the question.

The following criteria were created for extracting verbs with prefixes from song texts. Firstly, inseparable as well as separable verbs were extracted e.g. an inseparable verb *erzählst* in 16a or a separable verb *weitergehen* in 16b.

(16)   a. Von deinem Job **erzählst** du gar nichts oder wenig[30]

   b. Das ist ein Unding, so kann das nicht **weitergehen**[31]

Secondly, only separable verbs that were written together with their prefixes were selected — it was claimed that the user would benefit more from perceiving a simple verb[32] and its prefix as one unit. Separable verbs and their prefixes are written together, for instance, when the verb is an infinitive, as in example 16b, when the verb is used in a subordinate clause such as *rumliegt* in 17a; or when the verb is used in a form of participle e.g. *abgehört* in 17b.

(17)   a. [Und es tut weh, dass man sich nur sieht] Weil bei mir so viel Zeug von dir **rumliegt**[33]

   b. Die NSA hat seit Jahrzehnten jeden **abgehört**[34]

---

[29]AnnenMayKantereit: Oft gefragt
[30]Adam Angst: Ja ja, ich weiß
[31]Adam Angst: Professoren
[32]In "Hammer's German Grammar and Usage", the term *simple verb* means a verb without a prefix (Durrell, 2017)
[33]AnnenMayKantereit: Pocahontas
[34]Adam Angst: Splitter von granaten

The third criterion was also connected to separable verbs. Infinitives with *zu* derived from such verbs (e.g. *mitzuteilen* in the example 18) were not extracted, as *zu* would distract attention from the original prefix and might even trick the user into perceiving the original prefix and *zu* as one prefix.

(18)   Es besteht keine Not einem jedem davon **mitzuteilen**[35]

Fourthly, in case of past participles used in combination with an auxiliary (in passive voice, in perfect tense or in pluperfect tense), only the participle itself was extracted, such as *totgeschlagen* in 19a or *verändert* in 19b. Thus, in highlight mode only these parts of verb forms are highlighted. That helps to fix the attention on prefixes of verbs, not on their conjugation.

(19)   a.   Werden sie **totgeschlagen**, wenn sie kein Kopftuch tragen[36]
       b.   Denn was hat sich **verändert** in den letzten fünf Jahren?[37]

5.1.4 Passive voice

In contrast to spoken language, written language usually presents more challenges for an L2 learner of German. To express thoughts through writing accurately, certain skills are necessary, one of which is the usage of the passive voice (Koller et al., 2013). That skill can also be trained with interactive exercises in the app.

For the app, passive constructions in the past tense (*wurde geweckt* in 20a) and present tense (*wird gemacht* in 20b) were extracted. Constructions in the perfect tense were avoided due to their complexity, as illustrated by example 20c.

(20)   a.   Ich **wurde** heute morgen von 'nem Panzer **geweckt**[38]
       b.   Da **wird** ja schließlich nichts **gemacht**, außer viel Strom verbraucht[39]
       c.   [Hab Freunde betrogen und 'n paar auch verloren] Frauen verarscht und **bin verarscht worden**[40]

To sum all, four types of grammatical constructions were chosen for exercises in the app: *Prepositions and their cases*, *Verb conjugation*, *Verb prefixes* and *Passive voice*. All types of constructions had a number of criteria they were selected on.

---

[35]Audio88: Halleluija
[36]Adam Angst: Professoren
[37]Adam Angst: Splitter von granaten
[38]Alligatoah: Lass liegen
[39]Adam Angst: Splitter von granaten
[40]Lance Butters: 30

5.2 NLP Analysis Needed to Extract the Grammatical Constructions

In order to extract the chosen grammatical constructions from song texts, we needed NLP to analyze the language. Of particular relevance for us were the parts of speech of the words and their dependency relations. Here we, therefore, focus on dependency analysis and the different NLP tools available to conduct such analysis. Four dependency parsers trained on German texts were analyzed in this section: ParZu, spaCy, the Stanford Parser and the Mate tools parser.

5.2.1 Dependency parsers

ParZu (Sennrich et al., 2009) was trained on Europarl and includes the following components in its pipeline: sentence segmentation, tokenization, part-of-speech (POS) tagging, morphological analysis and the core component — dependency parser, with a preprocessing step before and a postprocessing step after it. Remarkably, for all the tasks except dependency parsing, ready-made tools were chosen e.g. the punkt_tokenizer from NTLK was used for tokenization, and clevertagger[41] — for POS-tagging. The POS-tagger uses the Stuttgart-Tübingen Tag Set (STTS). The dependency analysis is represented in the CoNNL dependency format and includes an index, a lemma, a POS-tag, a language-specific POS-tag, morphological features, a head and a dependency relation for every token. The three other parsers use a similar set of features.

spaCy[42] was trained on TIGER Corpus and WikiNER dataset, and it provides two models for parsing German texts: *de_core_news_sm* (the small model) and *de_core_news_md* (the medium model), with *de_core_news_md* performing with a higher accuracy. Linguistic features supported by spaCy include sentence segmentation, tokenization, POS-tagging, dependency parsing, and named entity recognition. Unlike other parsers, it does not assign morphological features to tokens in languages other than English. Both the POS-tagger and the dependency parser use the TIGER Treebank annotation scheme.

The Stanford Parser (Qi et al., 2018), which was trained on the Negra corpus, offers a Python package that makes the software easy to install and to use. The neural pipeline of the parser in the Python package includes tokenization, multi-word expansion, lemmatization, POS-tagging, morphological features tagging,

---

[41]https://github.com/rsennrich/clevertagger
[42]https://spacy.io/

and dependency parsing. The TIGER variant of STTS is used for POS-tagging, and grammatical relations are defined according to the Universal Dependencies representation[43].

The Mate tools provide a transition-based dependency parser with a joint POS-tagging and labeled dependency parsing (Bohnet, 2012). The German models, therefore, include a lemmatizer model and a joint parsing model consisting of a tagger, a morphological analyzer and a parser. Before applying the Mate tools to a text, it should be transformed into one-word per line 2009 CoNNL format[44] and tokenized. Although the Mate tools provide the script for the first transformation, they do not include a tokenizer. It is suggested to use an OpenNLP library[45] for that.

To test how well the parsers analyze lyrics, a song text of *Splitter von Granaten* by Adam Angst was used. It was randomly chosen from the lyrics containing all the four types of grammatical constructions.

Before comparing the analyses of the four parsers, the best spaCy model had to be chosen. Texts of songs are quite different from newspaper texts or Wikipedia articles that the models were trained on, so in order to make the best choice, the analyses of two models were compared.

Before applying a parser to the text, the latter was transformed in the way that every line of lyrics was represented as a separate sentence — either by preserving the punctuation mark in the end of the line or by inserting a full stop. Despite that, the small model split the sentence in the example 21 into two parts: *Keine Nachbarn* and *Nachts über Grenzen fliehen*. The medium model perceives the sentence as a whole.

(21)   Keine Nachbarn Nachts über Grenzen fliehen[46]

There was a number of cases when analysis of the small and the medium models differ. The analysis of punctuation marks was not taken into account, as it was not relevant for extracting grammatical constructions. Not including punctuation marks and tokens from the sentence in the example 21, which was split differently by the two models, thirty-nine cases were detected. The medium model was preferred, as it correctly analyzed twenty-one tokens, compared to thirteen by the small model.

---

[43]http://universaldependencies.org/docsv1/
[44]http://ufal.mff.cuni.cz/conll2009-st/task-description.html
[45]https://opennlp.apache.org/
[46]All the examples in section 5.2 were taken from song *Splitter von Granaten* by Adam Angst

The analyses of the four parsers exhibit some general differences and similarities, not necessarily connected to a certain grammatical construction. For instance, as has already been mentioned above, spaCy is the only parser that does not provide tokens with morphological characteristics. Morphological analysis of the Stanford Parser is easier to read than the analysis of ParZu and the Mate tool, as names of features are specified e.g. *Case=Acc | Gender=Masc | Number=Sing* for *Applaus* in 22.

(22)   Und die Welt spendet **Applaus**

Compared to spaCy, ParZu has more specific dependency relations inside the noun phrase — many tokens that are analyzed as *nk* (noun kernel) by spaCy have different relations in the analysis by ParZu e.g. *det* (determiner), *attr* (attributive) or *pn* (preposition complement). Dependency relations assigned by the Stanford Parser, however, are even more specific than the ones by ParZu. For instance, the relation of adjective to noun is *amod* (adjectival modifier), and of cardinal number to noun — *nummod* (numeric modifier), while both of the relations are marked *attr* in the analysis of ParZu.

The tokenizer by the OpenNLP library, which is recommended by the Mate tools, does not behave completely usual, compared to tokenizers of other parsers — for example, it treats two sentences with the question mark after the first sentence as one sentence. Moreover, it splits the line illustrated in example 23 in two sentences, with the first sentence ending with an ellipsis. ParZu, spaCy and the Stanford Parser treat this line as one sentence.

(23)   Asylbewerberheime sind doch sicher, alles klar... 43 Anschläge und dass in einem Jahr

In the next subsections, more specific differences between the analyses of the four parsers are discussed.

5.2.2 Prepositions and their cases

The Stanford parser is the only parser that transforms a contraction of a preposition with the definite article e.g. *am* or *aufs* in two separate words in its analysis. If a combination is non-standard, however, it is treated as an adverb. That makes it impossible to extract constructions like *zum Vergnügen* or *unterm Tellerrand* using the Stanford parser.

In contrast to ParZu, no case is assigned to prepositions in the Stanford Parser and the Mate tools. Moreover, the Mate tools, unlike the three other parsers, assigns

some of the tokens wrong morphological characteristics. Therefore, some of the future rules for the Mate tools would not be able to include morphological features of tokens, as it would lead to missing a few constructions.

ParZu seems to be the best tool for extracting prepositional phrases, as its analyses are right, compared to the Mate tools, and it is able to extract constructions like *zum Vergnügen* and *unterm Tellerrand*, in contrast to the Stanford parser. Apart from that, unlike spaCy, ParZu includes morphological information, which helps to describe prepositional phrases that need to be extracted more precisely.

5.2.3 Verb conjugation

First of all, finite verb forms should have right language-specific POS-tags, so that they could be extracted. However, in order to generate distractors for such verb forms, their morphological characteristics and their infinitives are also needed. Therefore, all these three features should be taken into account while choosing a parser. spaCy does not provide morphological information, so it is not suitable for the current topic. Nevertheless, it was compared to other parsers to show how well it had dealt with other aspects of grammar analysis.

There is no dependency parser that analyzed all the finite verb forms in the right way. With twenty-eight verb forms to extract in total, spaCy assigned wrong language-specific part-of-speech tags to six of them — for instance, *provoziert* in 24a was tagged *VVPP* (perfect participle, full) instead of *VVFIN* (finite verb, full). Other examples included cases when a verb was tagged *VVFIN* instead of *VVPP* and *VVINF* (infinitive, full) instead of *VVFIN*. Two of the verbs had wrong infinitives — for example, *taten* in 24b was assigned infinitive *taten* instead of *tun*. In total, eight out of twenty-eight verbs were analyzed wrong by spaCy.

(24) a. Das 'nen Atomkrieg **provoziert** und denkt es wäre 'ne Kissenschlacht

     b. Und wir **taten** überrascht und waren 'ne Woche lang empört

In case of ParZu, five verbs were assigned wrong language-specific POS-tags. As in the analysis by spaCy, some of them were tagged *VVINF* instead of *VVFIN* and some — *VVPP* instead of *VVFIN*. ParZu did not provide any wrong morphological analysis, but for a few verb forms, some morphological features were not determined e.g. *geht* in 25a was not assigned neither number nor person. Cases when only the mode was not determined were not considered a mistake — for instance, *wissen* from example 25a was not assigned any mode. From analyses of a greater

number of songs with the parser, it could be concluded that for every such case the indicative mode could be assumed. Thus, including the verb *abgehört* in 25b that was assigned a wrong infinitive *abgehören*, ten out of twenty-eight verbs were analyzed wrong by ParZu.

(25) a. Doch worum es gerade **geht** wissen wir selbst nicht so genau
      b. Die NSA hat seit Jahrzehnten jeden **abgehört**

In the analysis by the Stanford parser, the following mistakes could be found. Two of the verbs were assigned wrong language-specific characteristics and got tagged *VVINF* instead of *VVFIN*. This mistake occured in the analyses of all parsers, including the Mate tools. For example, *springen* in 26a was wrongly assigned tag *VVINF* by the Stanford parser, ParZu and spaCy. One other verb, *taten* in 26b, which also posed challenges to spaCy, was assigned wrong morphological features and a wrong infinitive by the Stanford parser. All in all, analyses of only three out of twenty-eight verb forms contained mistakes in case of the Stanford parser.

(26) a. [Solang hier keine Sirenen erklingen] Keine Soldaten durch unsere Fenster **springen**
      b. Und wir **taten** überrascht und waren 'ne Woche lang empört

Analysis by the Mate tools contains a number of various mistakes. One of the verbs was assigned wrong morphological features, another one — a wrong infinitive, while the third one, *schießen* in 25a had a combination of these mistakes. Two other verbs were not assigned any morphological characteristics at all. Finally, language-specific POS-tag for one of the verbs was selected wrong. Therefore, six out of twenty-eight verbs were assigned wrong analysis by the Mate tools.

(27) Polizisten ticken aus und **schießen** wahllos in die Menge

As has been mentioned above, not only POS tags, but also morphological features of verbs were important for extracting finite verb forms. Therefore, taking into account the results of the parsers that provide morphological information about word forms, the best parser for extracting finite verb forms was the Stanford parser.

5.2.4 Verbs with prefixes

For getting a list of verbs with prefixes from a song text, candidates for such verbs should be identified first. That is exactly why the analysis from a dependency

parser is needed — to get all possible verb forms that then will be tested for having a prefix.

It is worth mentioning that the extraction of verbs with prefixes is rather similar to the extraction of finite verb forms. One of the differences is including forms in passive voice for verbs with prefixes and excluding them for verb conjugation, as verb forms in passive voice is a grammar topic on its own. Another difference is infinitives, which can serve as candidates for verbs with prefixes, but do not constitute a finite verb form. Moreover, in contrast to the topic *Verb conjugation*, infinitives and morphological characteristics are not needed for generating distractors for the current topic. Therefore, it was decided to analyse the results of the parsers for the two topics separately.

As neither morphological features nor infinitives were taken into account, the main criterion for choosing the most suitable parser was its ability to assign right language-specific POS-tags.

In the analyses of spaCy, ParZu and the Mate tools, some verb forms were assigned wrong language-specific POS-tags — for example, *erklingen* in 28a was assigned tag *VVINF* instead of *VVFIN* by ParZu or *abgehört* in 28b was tagged *VVFIN* instead of *VVPP* by spaCy. In the first example, it was rather clear that the wrong tag was assigned due to the ambiguity of the verb form. In the second case, however, the verb forms were not homonymous — there was no finite verb form *abgehört*. A mistake of the second type was only found in the analysis of spaCy. ParZu and the Mate tools analyzed verb forms in a wrong way only if the forms had homonyms. There were three mistakes in assigning language-specific POS-tags in the analysis of spaCy, while both ParZu and the Mate tools each produced two such mistakes.

(28)   a.   Solang hier keine Sirenen **erklingen**

        b.   Die NSA hat seit Jahrzehnten jeden **abgehört**

In contrast to the parsers discussed above, the Stanford Parser assigned all verb forms correct language-specific POS-tags. For this reason, it was considered the best parser for extracting verbs with prefixes.

5.2.5 Passive voice

ParZu, spaCy and the Mate tools did not distinguish perfect tense and passive voice — namely, the dependency relation between an auxiliary and the main verb was

always *aux* (auxiliary verb; in case of ParZu) or *oc* (clausal object; in case of spaCy and the Mate tools). Thus, constructions like *hat verändert* in 29a and *wird gemacht* in 29b were parsed alike.

(29)   a.   Denn was **hat** sich **verändert** in den letzten fünf Jahren
       b.   Da **wird** ja schließlich nichts **gemacht**, außer viel Strom verbraucht

The Stanford parser, however, had a special relation for verb forms in the passive voice — *aux:pass*, while for the perfect tense the relation was simply called *aux*. Therefore, the Stanford Parser was best suitable for extracting constructions for this topic, as well.

To make the conclusion, ParZu will be used for extracting prepositional phrases, and the Stanford Parser — for extracting constructions for the rest of the topics. However, on the stage of writing rules it will become clear that for some topics, the results of the analysis of one song can not be generalized for the whole corpus of song texts. For these topics, a parser will be chosen again, this time based on a greater selection of songs.

5.3 Manual Tagging and Splitting

To be able to write rules for extracting the grammatical constructions discussed in section 5.1 and check the rules afterwards, all the four construction types — prepositions and their cases, verb conjugation, verbs with prefixes and passive voice — were manually tagged in all the song texts. During that process, the limitations described in section 5.1 for every topic were met.

After that, songs containing constructions for each type were grouped together. The distribution of constructions of different types can be seen in table 5.1.

Prep. and their cases	Verb conj.	Verbs with pref.	Pass. voice
130	131	121	35

Table 5.1: Number of songs with constructions of each type

One could see that finite verb forms and prepositional phrases could be found in all or almost all constructions, while verbs with prefixes occurred in 92% of songs. Verb constructions in the passive voice were, however, relatively rare and could be seen in only 27% of texts.

For each construction type, songs were split into two sets — one set was used for writing rules for extracting a certain type of constructions, and the other — for

evaluating the rules. Here, these sets are called with terms often used in machine learning — *training set* and *test set*. The ratio 20:80 was chosen for the split, following the Pareto principle[47].

It was decided to sample every set with respect to the genre of songs, in order to account for possible differences in the structure of constructions or in their rate of occurrence among various genres. As stated in section 4.1, participants assigned genres to artists and bands based on their own definition of genres. In two cases, it led to a conflict of genres e.g. some students labeled Rammstein as *Rock*, and other — as *Metal*. In such cases, the genre was assigned based on the number of people that had chosen every genre. Rammstein was assigned genre *Rock*, as it was the option preferred by most students.

	Total N of songs	Training set	Test set
Dance/Electronic/House	3	1	2
Hip-Hop/Rap/Trap	42	8	34
Metal	3	1	2
Pop	30	6	24
Rock	38	8	30
Singer/Songwriter	14	3	11
Total	130	27	103

Table 5.2: Distribution of songs by genre in training and test sets for construction type *Prepositions and their cases*

Sampling the training and the test sets by genre for construction type *Prepositions and their cases* can be illustrated by table 5.2. In the same way, the training and the test sets for other construction types were balanced on music genres.

5.4 Rules

5.4.1 Prepositions and their cases

As has been explained in section 5.2, ParZu was chosen as the best suitable parser for extracting prepositional phrases from song texts. Before applying the parser to the texts, they were prepared in the same way as the song text *Splitter von Granaten* by Adam Angst that was used for choosing a parser for each type of constructions. Namely, to every lyrics line that did not have a punctuation mark at the end was

---

[47]https://en.wikipedia.org/wiki/Pareto_principle

attached a full stop. Then all lines were written as a prose text, being separated with a space.

One could argue that representing each line of song lyrics as a separate sentence may lead to errors in parsing. However, only 7% of the lines in all the song texts contain a punctuation mark signaling the end of the sentence at the end of the line. For many songs e.g. *Lieder* by Adel Tawil, a comma is the only punctuation mark found in the text. Representing a whole song text as one sentence would, undoubtedly, lead to many more errors than the suggested representation of the song text.

For representing words analyzed by ParZu, a class *Word* was written. The class includes all the attributes needed to describe the information the parser provides about a word form: *id* for the index of the word in the sentence, *form* for the word form itself, *lemma* for the lemma of the word form, *upos* for its POS-tag, *xpos* for a language-specific POS-tag, *head* for the id of the word form's head and *rel* for the type of dependency relation of the word to its head. Moreover, class *Word* has an attribute *feats* — a Python dictionary that contains word form's morphological features. Depending on the part of speech, this dictionary can contain different keys. For instance, if the word is an article, the dictionary would only contain its case, and if the word belongs to finite verbs — its person, number, tense and mode. Class *Word* was also used for representing words that constituted rules.

Every rule was represented as an array of strings, similar to the array in figure 5.1. Strings in a rule got subsequently transformed into instances of Word class — for this reason, all attributes in strings were separated with a tab, as in words analyzed by ParZu. A string had to include a language-specific POS-tag and could include restrictions on morphological features, the head of the word form and the dependency relation of the word form to its head.

```
[
  "APPR\t\t\t",
  "PPOSAT\t_|0|_\t2\tdet",
  "NN\t1|0|1\t0\tpn"
],
```

Figure 5.1: Example of a rule for extracting prepositional phrases

Language-specific POS-tags and types of dependency relations were the same as in analysis of ParZu e.g. *PPOSAT* (attributive possessive pronoun) and *det* (determiner). Morphological features, however, instead of looking more specific,

such as *Neut | Acc | Pl* (neuter gender, accusative case, plural number), showed the agreement between the words constituting the construction. For example, as shown in figure 5.1, *PPOSAT* and *NN* (noun) each have three morphological features separated with a vertical bar — gender, case and number. Both the pronoun and the noun should have the same case as the first element of array (it has index 0), tagged as *APPR* — the preposition. There are no limitations set on the gender and the number of the pronoun. However, the noun's gender and number should agree with the gender and number of the second element of array (it has index 1), namely, the pronoun. Heads of words in a rule were also specified relative to the structure of the array. Thus, the possessive pronoun has the third element of array (it has index 2) as its head — the noun, and the head of the noun is the first element of the array, the preposition.

Apart from containing strings with word's characteristics, as described above, rules could contain a special character +, meaning that the word described by the previous string could be repeated an unlimited amount of times. Figure 5.2 gives an example of such rule, describing a prepositional phrase that includes a preposition, a possessive pronoun, one or more adjectives and a noun.

```
[
  "APPR\t\t\t",
  "PPOSAT\t_|0|_\t-1\tdet",
  "ADJA\t_|1|0|1|_\t-1\tattr",
  "+",
  "NN\t1|0|1\t0\tpn"
],
```

Figure 5.2: Example of a rule with the special character +

To exclude erroneous constructions, it might be useful to set restrictions on the lemmas that in some cases lead to false positives. For instance, *bis* in 30 is analyzed as a preposition, which leads to extracting *bis der Tod* as a prepositional phrase. A blacklist was introduced as means of excluding constructions with some lemmas from the result. The blacklist is a dictionary, where lists of excluded lemmas correspond to the language-specific parts of speech.

(30)   Willst du **bis** der Tod euch scheidet [Treu ihr sein für alle Tage?][48]

There was also a way to specify word forms or parts of speech that should not

---

[48]Rammstein: Du hast

precede or follow certain constructions. For example, extracting *Zum Homo* from the line in example 31 is wrong, as the whole prepositional phrase also includes the word *Sapiens*, tagged as NE (named entity). Thus, a rule was introduced that a noun in the construction could not be followed by a named entity. Another dictionary, representing a blacklist for words surrounding the constructions, contained such rules and helped to avoid cases similar to the one described above.

(31)   **Zum Homo** Sapiens gehört nicht nur aufrecht geh'n[49]

In total, there were 391 prepositional phrases in the training set. All morphological features of the word forms from 88.5% of the constructions were defined, but 11.5% of the constructions lacked some characteristics. For example, no gender and number was assigned to the possessive pronoun and the noun in *auf deine Fragen* (see example 32a). The single construction for which a morphological characteristic was defined wrong was also included in these 11.5% — *Bett* was assigned nominative case in the prepositional phrase *im Bett* (the lyrics line containing the construction is presented in example 32b).

(32)   a.   Kein Bock **auf deine Fragen**[50]

   b.   Danach, bei dir, du nackt **im Bett** und ich barfuß am Klavier[51]

Therefore, two sets of rules were defined: a set of strict rules that were not able to extract these 11.5% of constructions from the training set and a set of non-strict rules that accounted for the constructions that lacked morphological features. An example of a strict rule for extracting a prepositional phrase consisting of a preposition, a possessive pronoun and a noun is presented in figure 5.3, while the same rule in a non-strict version is given in figure 5.4. This non-strict rule could be used for extracting constructions like *auf deine Fragen* from example 32a.

```
[
  "APPR\t\t\t",
  "PPOSAT\t_|0|_\t2\tdet",
  "NN\t1|0|1\t0\tpn"
],
```

Figure 5.3: Example of a strict rule

[49]Farin Urlaub: IDisco
[50]Bratze: Die auswendigen Muster
[51]AnnenMayKantereit: Barfuß am Klavier

As could be expected, the strict rules led to 45 false negatives which constituted the 11.5% of the constructions that lacked morphological characteristics. Non-strict rules yielded not only some false negatives, but also a few false positives.

```
[
  "APPR\t\t\t",
  "PPOSAT\t_|0|_\t2\tdet",
  "NN\t_|0|_\t0\tpn"
],
```

Figure 5.4: Example of a non-strict rule

Maintaining high precision was considered more important than achieving a high recall, because providing the user with less correct constructions is definitely more effective for acquiring a certain grammatical topic than getting more correct constructions on the cost of selecting constructions completely not suitable for the chosen grammatical topic. The second case might frustrate the user, and even incline him/her to memorize non-existing rules. Thus, avoiding false positives had a higher priority than avoiding false negatives, and, if needed, was even done on the cost of getting more false negatives.

For this reason, the non-strict rules that led to false positives were modified. Some of them were made stricter e.g. the rule for extracting a prepositional phrase consisting of a preposition, an article and a noun. This rule accidentally extracted construction *für ein Hurensohn* from the line presented in example 33a and was, therefore, modified by making all word forms in a prepositional phrase share the same case. Some of the non-strict rules were deleted, too — for instance, the rule used to extract *von der Geschicht'* from line in example 33b also extracted *nach dem Ende '* from the sentence in 33c, while the right prepositional phrase *nach dem Ende* did not include the apostrophe. After these two strategies were applied to modifying the rules, all the false positives extracted by the set of the non-strict rules could be avoided.

(33) a. Was bist du **für ein Hurensohn**, sogar die Wurst hat zwei![52]

b. Und die Moral **von der Geschicht'**: "Es heißt Geschichte!"[53]

c. Ich bin raus, kann schon **nach dem Ende '**nen Anfang sehen[54]

---

[52]Audio88: Gnade
[53]Audio88: Gnade
[54]Casper: Auf und davon

As the non-strict set of rules led to extracting more constructions from song texts, even after some of its rules were modified, it also showed better results on the training set than the strict set. All in all, 20 constructions could not be extracted by the non-strict rules.

The black list was a part of both strict and non-strict sets of rules. One of the members is *bis*, which cannot be a preposition, as it leads to extracting *bis der Tod* from example 30, repeated here in example 34a. The other member is *paar*, which, unlike other attributive indefinite pronouns with a determiner[55], is unable to illustrate the case required by a preposition in the context represented in example 34b.

(34)   a. Willst du **bis** der Tod euch scheidet [Treu ihr sein für alle Tage?][56]

         b. Ich hock in meinem Viertel **mit paar Homies** auf 'ner Bank[57]

Some constructions were written incorrectly in song texts, which hindered the parser from extracting them — for instance, *am Gaffen* in example 35 was written using a small letter — *am gaffen*. Such mistakes were corrected in texts, and the corrected texts were analyzed by the parser again.

(35)   Viel zu viele Menschen um uns rum sind **am Gaffen**[58]

As the set of the non-strict rules showed better results on the training set, we took it as a basis for a final set of rules. However, before being applied to the test set, the rules were made more general, so that they could account for some constructions that were not found in the training set. A good rule should be as general as possible — however, without overgeneralizing and leading to extracting false positives.

First of all, every rule for extracting constructions with adjectives was modified with the special character + to include more than one adjective. For instance, the rule shown in figure 5.5 originally contained no + and was used for extracting constructions like *Nach all den einsamen Jahren* in example 36. After the rule was modified, an unlimited number of adjectives could be extracted as part of the prepositional phrase described by the rule e.g. *Nach all den einsamen traurigen langen ... Jahren*.

---

[55]Translation of "attribuierendes Indefinitpronomen mit Determiner", term used in the guidlines for tagging texts with the tagset of STTS (Schiller et al., 1999). This tagset is used by a POS-tagger in ParZu.

[56]Rammstein: Du hast

[57]Haze: Becher & Blunt

[58]Juju: Hardcore high

```
[
  "APPR\t\t\t",
  "PIDAT\t_|0|_\t-1\tdet",
  "ART\t_|1|0|1\t-1\tdet",
  "ADJA\t_|1|0|1|_\t-1\tattr",
  "+",
  "NN\t1|0|1\t0\tpn"
],
```

Figure 5.5: Example of a rule modified with the special character +

(36) **Nach all den einsamen Jahren** [Macht alles einen Sinn und der bist du!][59]

The only rule that was not modified to include more than one adjective is shown in figure 5.6. This rule extracts constructions like *Mit müdem Lächeln* in example 37, where an adjective serves as a word that illustrates the case the preposition requires. Therefore, if at least one more adjective is used after the first adjective, it would be too easy for the user to choose the right option — simply by comparing all the possible options against the second adjective he/she would arrive at the right answer.

(37) Bringe ich **mit müdem Lächeln** endlich Frieden für die Szene

```
[
  "APPR\t\t\t",
  "ADJA\t_|_|_|_|_\t-1\tattr",
  "NN\t_|_|_\t0\tpn"
],
```

Figure 5.6: Example of rule not modified with a special character +

However, not every rule had an equivalent for extracting the same construction with an adjective. Not taking into account two rules containing personal pronouns that cannot be used with adjectives, seven out of nineteen rules did not have such equivalents. To get proof that constructions which would be extracted with these equivalents were possible, the TIGER corpus[60] was used.

---

[59]Helene Fischer: Achterbahn
[60]https://www.ims.uni-stuttgart.de/en/research/resources/corpora/tiger/

```
[
  "APPR\t\t\t",
  "PIAT\t_|0|_\t2\tdet",
  "NN\t1|0|1\t0\tpn"
],
```

Figure 5.7: Example of rule that did not have an equivalent rule including an adjective

In the corpus, constructions corresponding to the equivalent rules for four out of seven rules were found — for instance, for the rule shown in 5.7 and used for extracting such constructions as *in jedem Reim* in example 38.

(38)   Nehm das Mic denn mein Seelenheil steckt **in jedem Reim**[61]

The equivalent rule containing an adjective is presented in figure 5.8. The following construction from the corpora corresponded to the rule — *zu vielen britischen Politikern* in example 39.

(39)   Anders als sein Vorgänger Lord Wilson, der in der Tradition beamterer Sinologen aus London "Foreign Office" stand, verfügt Chris Patten über direkten Zugang **zu vielen britischen Politikern**.

```
[
  "APPR\t\t\t",
  "PIAT\t_|0|_\t-1\tdet",
  "ADJA\t_|1|0|1|_\t-1\tattr",
  "+",
  "NN\t1|0|1\t0\tpn"
],
```

Figure 5.8: Example of rule equivalent to 5.7 and containing an adjective

Generalized rules gave the same result on the training set as the set of the non-strict rules, as generalizations were aimed at constructions not found in the training set.

After the set of the non-strict rules was generalized using the two strategies discussed above, it was applied to the test set containing 1488 prepositional phrases. In the results, different types of mistakes could be found: false positives, false

---

[61]Haze: Becher & Blunt

negatives and constructions selected in a wrong way. To avoid the false positives and the constructions selected in a wrong way, some rule modifications were necessary. A few new rules were written e.g. the one illustrated in figure 5.9. Without it, instead of the construction *im gegenüber liegenden Gebäude*, a construction *gegenüber liegenden Gebäude* was extracted from line in example 40.

(40)   Doch die Kripo sitzt **im gegenüber liegenden Gebäude**[62]

Apart from that, the blacklist, as well as the blacklist for words surrounding the constructions acquired some new members.

```
[
  "APPRART\t\t\t",
  "APPR\t0\t\t",
  "ADJA\t_|-1|0|-1|_\t-1\tattr",
  "NN\t_|0|_\t1\tpn"
],
```

Figure 5.9: Example of rule written after analyzing results on test set

In this way, we managed to avoid all false positives and constructions extracted in a wrong way, which, however, cost some more false negatives. The final variant of the rules led to 179 false negatives.

Applying the final rules to the training set also led to some more false negatives — three more constructions could not be extracted with the new set of rules.

Set	Precision	Recall
Training set	100%	94.1%
Test set	100%	88%
Total	100%	89.2%

Table 5.3: Precision and recall for extracting constructions of type *Prepositions and their cases*

The final results could be seen in table 5.3. In both training and test set, the precision equals to 100%. The recall, however, is not perfect. The recall of extracting constructions from the training set equals to 94.1%, which is 6.1% higher than the result on the test set. The total recall amounts to 89.2%.

---

[62]Haze: Weisch Weisch

5.4.2 Verb conjugation

While creating rules for extracting verb forms with the analysis by the Stanford parser, it was stated that there were more errors in the analysis than it had been expected based on the test analysis of *Splitter von Granaten* by Adam Angst. For this reason, it was decided to run all the parsers that include morphological analysis on the training set for finite verb forms, in order to get a more reliable estimation of the parsers' performance. In every analysis, identical lines of lyrics were eliminated, so that the comparison of the parsers would be less time-consuming.

The results can be seen in table 5.4. ParZu provided the greatest number of correct analyses of verb forms, while the Mate tools analyzed 15.7% less verb forms in a right way. The Stanford parser, initially chosen for extracting finite verb forms, showed the worst result.

Parser	Analyzed right
ParZu	**76.8%**
Stanford parser	59.1%
Mate tools	61.1%

Table 5.4: Percentage of the verb forms in the training set analysed correctly by each parser. Topic *Verb conjugation*

In table 5.5, the distribution of mistakes made by each parser is illustrated. A verb form analyzed in a wrong way could contain several types of mistakes at the same time.

Parser	Morph. n.d.	Infinitive	Morph. wrong	POS	Dep. rel.
ParZu	**66.4%**	45.7%	27.6%	25%	0%
Stanford	2.9%	**78.4%**	67.2%	9.8%	0.5%
Mate tools	13.9%	**76.3%**	57.7%	24.2%	0%

Table 5.5: Distribution of errors for the analysis of the training set by each parser. Topic *Verb conjugation*

As one could see, the most common error of ParZu was not defining morphological features of a verb form. While not being assigned a mode was not considered a mistake, *denke* in 41a was also not assigned a number.

(41)   a.   [Ich lieb die Kühle und das fahle Licht] Und **denke** schon fast gar nicht
            mehr an dich[63]

---

[63]Isolation Berlin: Aquarium

b. Ich **empfing** diese Worte aus den Wolken – Amen![64]

c. Ich **ging** wie ein Ägypter[65]

d. **Hör** auf, dich zu wehren, das macht doch keinen Sinn[66]

e. Dort wo wir gefahren **sind**[67]

For the Stanford parser and the Mate tools, the most frequent error was determining an infinitive of a verb in a wrong way. For example, the Stanford parser defined the infinitive of *empfing* in example 41b as *empfehlen* instead of *empfangen*. Assigning wrong morphological characteristics could be demonstrated on *ging* in example 41c. The Mate tools analyzed it as a verb in the present tense, while the form belongs to the past tense. Example 41d illustrates assigning a wrong POS, as ParZu analyzed the verb form *Hör* as a noun. The Stanford parser was the only parser that had a mistake of type *wrong dependency relation* — it assigned the dependency relation *aux:pass* to *sind* in example 41e, which would prevent extracting *gefahren sind* as a verb form in the perfect tense.

```
[
  "VAFIN\t?|?|Pres|Ind\t\t",
  "VVPP\t\t0\taux"
],
```

Figure 5.10: Example of a rule for extracting finite verb forms

As ParZu showed the best result on the training set, it was chosen as a suitable parser for extracting finite verb forms instead of the Stanford parser that had been chosen for this purpose initially. The set of rules for extracting verb forms had the same structure as the set of rules for extracting prepositional phrases. For instance, the rule shown in figure 5.10 was used for extracting a verb form in the perfect tense, consisting of an auxiliary verb and a past participle, such as *hat gehört* in example 42. The sign *?* stands for any possible value.

(42)   Wir waren jung und unbeschwert, die ganze Welt **hat** uns **gehört**[68]

The blacklist for extracting finite verb forms contained *werden*, so that such rules as the rule in figure 5.10 could not extract verb forms in passive voice with *werden* as

---

[64]Audio88: Halleluja
[65]Adel Tawil: Lieder
[66]Mark Forster: Chöre
[67]Frittenbude: Einfach nicht leicht
[68]Andrea Berg: Ja, ich will

an auxiliary verb, and *zeihen*. In a strange way, any form of verb *ziehen* was assigned infinitive *zeihen* by ParZu, which would hinder generating suitable distractors for the verb form. The blacklist for surrounding words helped to avoid verb forms exclusively used in spoken German, such as *Tanz'* in example 43, by excluding all verb forms that had an apostrophe written after them. The set of rules remained the same for the training and the test set.

(43)   [Und ich singe diese Lieder] **Tanz'** mit Tränen in den Augen[69]

However, there were a number of issues leading to false positives in both sets. Firstly, some infinitives were analyzed as finite forms by ParZu and were, therefore, extracted. For this reason, the analyses of the training set by the three parsers were compared again in order to determine which parser produced the best result in distinguishing the homonymous forms and could be used to reduce the number of false positives.

Parser	Analyzed right
ParZu	**93.5%**
Stanford	92.4%
Mate tools	91.8%

Table 5.6: Percentage of infinitives in the training set analyzed correctly by each parser

As could be seen from table 5.6, the other parsers provide even worse results than ParZu. Therefore, infinitives analyzed as finite forms could not be filtered out by the analysis of the Stanford parser or the Mate tools. They will be used for multiple choice questions exactly as the finite verb forms.

Secondly, verb forms in the passive voice with *sein* as an auxiliary e.g. *waren geboren* in example 44a were often extracted by the rules used for verb forms in the perfect or pluperfect tense, such as *war angelangt* in example 44b.

(44)   a.   Wir **waren geboren**, um zu leben für den einen Augenblick[70]

   b.   Ich **war** am Ende der Straße **angelangt**[71]

To avoid such cases, a list of infinitives that could be used with *sein* in the perfect tense was needed. As a source for the list, the database of German verbs created

---

[69]Adel Tawil: Lieder
[70]Unheilig: Geboren um zu leben
[71]Adel Tawil: Lieder

as part of the thesis project by Sabine Buchholz was used (Buchholz, 1996). In the database, every verb is mapped to an auxiliary which is used to build a perfect tense of this verb. The verbs that can combine with both *sein* and *haben* in the perfect tense e.g. *schwellen* were also included in the list.

It was, therefore, checked that the infinitive of every participle extracted with *sein* was found in the list, otherwise the selected construction was filtered out. Such filter helped to significantly reduce the amount of verb forms in the passive voice that had been extracted by the set of rules before. The only participle that still led to constructions extracted in a wrong way was *zerrissen*. Two false positives with this participle are shown in example 45.

(45)   a. Ich **bin** grade so krass **zerrissen**[72]

       b. Meine Hose **ist zerrissen** und der Wagen ist hin[73]

Apart from that, some text corrections were made to improve the result on the training and test set. Some of the mistakes were connected to forms in the imperative mode — for instance, an apostrophe could be written in a verb form used in standard German, such as *Guck'* in example 46a. In such cases, the apostrophe was deleted, as it hindered the extraction of the verb form. In other cases, no apostrophe was used, although the verb form was contracted e.g. in *Lächel* from the line in example 46b. As we wanted to avoid extracting the forms used exclusively in spoken German, an apostrophe was inserted after such forms.

(46)   a. **Guck'** mal, Jutta, da schwimmt unsre alte Mikrowelle[74]

       b. **Lächel** doch mal, ich geb' dir auch einen aus[75]

       c. [Immer wenn ich einsam bin] **Geh** ich ins Aquarium[76]

       d. [Doch sie sei leider pleite Hätt' von ihrem letzten Geld] Die Karte **heut** gekauft[77]

Apostrophes were also inserted after some verb forms of 1st person, singular number and present tense. For instance, *Geh* in example 46c was extracted as a verb form in the imperative mode, before the apostrophe was added after it. In a similar way, to prevent an adverb *heut* from being extracted as a verb, an apostrophe was

---

[72]Juju: Vermissen
[73]Cro: Whatever
[74]Alligatoah: Lass liegen
[75]Audio88: Schellen
[76]Isolation Berlin: Aquarium
[77]Frittenbude: Bilder mit Katze

inserted after it. Example 46d illustrates one of the contexts where it had to be done.

Other singular cases of text corrections include writing *werd'* as a full form to prevent the participle *verfolgt* from being extracted as a finite verb form (see example 47a) and writing infinitival nominals, such as *taktieren* in example 47b, with a capital letter.

(47)   a.   Doch irgendwann **werd'** ich vom Glück verfolgt (mmh)[78]

   b.   Kein **taktieren**: anvisieren, konfrontieren, triumphieren![79]

Although many strategies of reducing the number of false positives were applied, some of them were impossible to avoid. In the training set, the extraction of finite verb forms produced 177 false negatives and 5 false positives, and in the test set — 599 false negatives and 37 false positives. The results could be seen in table 5.7. Despite the number of false positives, precision remains rather high and amounts to 98.4% for the combination of the training and test set. The recall, however, is not greater than 77% for both sets, and equals 76.4% for their combination.

Set	Precision	Recall
Training set	99.1%	76.3%
Test set	98.1%	76.7%
Total	98.4%	76.4%

Table 5.7: Precision and recall for extracting constructions of type *Verb conjugation*

The types of false positives that could not be avoided are given in example 48 and illustrate the following types of misanalysis by ParZu: an infinitive analyzed as a finite form (48a), a past participle analyzed as a finite form (48b), a prefix of the verb analyzed as a finite form (48c), an adverb analyzed as a finite form (48d) and, lastly, a verb form in the subjunctive mode analyzed as a verb form in the past tense (48e).

(48)   a.   [Ich sehe oben am Himmel] Flügzeuge **verschwinden**[80]

   b.   Der Pudel **parfümiert**, das Lächeln ist aus Stahl[81]

   c.   Und es tut **weh** dich schon wieder so [Wieder zu sehen][82]

---

[78]Peter Fox: Haus am See
[79]Farin Urlaub: Herz? Verloren
[80]Max Giesinger: Legenden
[81]Marteria: Verstrahlt
[82]AnnenMayKantereit: Pocahontas

d. Des machen die doch **bestimmt** auch nicht so, bei sich in Afrika[83]

e. Ich **wünschte**, ich wäre gegen all die Pheromone immun[84]

5.4.3 Verbs with prefixes

Since it was stated that the analysis by the Stanford parser was not good enough for extracting constructions of type *Verb conjugation*, it was decided to re-evaluate the quality of different parsers for topics *Verbs with prefixes* and *Passive voice*, too. Similarly to the topic *Verb conjugation*, the re-evaluation was done using a training set.

The results for the topic *Verbs with prefixes* are presented in table 5.8. ParZu shows the best result by analyzing 79.1% of the verb forms correctly. The result of the Stanford parser is rather close to ParZu and differs by 5.6%. The analysis of the Mate tools, however, contains 12.1% more mistakes than the analysis of ParZu.

Parser	Analyzed right
ParZu	**79.1%**
Stanford parser	74.7%
Mate tools	67%

Table 5.8: Percentage of the verb forms in the training set analysed correctly by each parser. Topic *Verbs with prefixes*

The distribution of mistakes in the analysis of each parser is shown in table 5.9. In order to generate distractors for topic *Verb conjugation*, many features of a verb form should be determined right, such as morphological characteristics and the infinitive of the verb. Moreover, a part of speech, and, in some cases, a dependency relation are needed for extracting the constructions. A verb form extracted for topic *Verb prefixes*, however, should be assigned only a correct infinitive and a correct part of speech. Both errors could be sometimes found in a single verb form.

Parser	Infinitive	POS
ParZu	**68.4%**	**68.4%**
Stanford	56.5%	**78.3%**
Mate tools	**80%**	60%

Table 5.9: Distribution of errors for the analysis of the training set by each parser. Topic *Verbs with prefixes*

---

[83]Audio88: Schellen
[84]Farin Urlaub: Herz? Verloren

```
[
    "VVFIN\t\t\t"
],
```

Figure 5.11: Example of a rule for extracting verbs with prefixes

Analysis of ParZu contained the equal number of errors in determining infinitives and parts of speech. As for the Stanford parser, more errors in assigning a correct part of speech could be found. For instance, *vertraut* in example 49a was assigned POS-tag *ADJD* (adjective), instead of being analyzed as a verb. The Mate tools, on the other hand, encountered more difficulties with determining an infinitive of a verb form in a right way. Instead of *erzählen*, *erzählst* in example 49b was assigned an infinitive *erzählsen* by the Mate tools.

(49)   a. Ich will dass ihr mir **vertraut**[85]

       b. Alles, was du so **erzählst** [Hört sich irgendwie nice an][86]

Thus, a set of the rules for extracting verbs with prefixes was based on the analysis by ParZu and contained four very simple rules: for a finite verb (as in figure 5.11), an infinitive, a past participle and a verb in the imperative mode. No morphological characteristics, heads or dependency relations had to be specified. The blacklist did not have any members, while the blacklist for surrounding words was used for excluding all verb forms with an apostrophe written after them, as they were not used in standard German. A verb *überleb'* in example 50 serves an illustration for such verb form.

(50)   [Alle war'n sich einig] Dass ich niemals **überleb'**[87]

The set of rules did not change from the training to the test set. In the results on the training set, 22 false negatives and no false positives were found, while the results on the test set included both false negatives and false positives — 141 and 4 instances correspondingly. Precision and recall could be seen in table 5.10. As the set of rules produced no false positives on the training set, precision equals 100%. For the test set and a combination of the training and the test sets, it is also close to perfect and equals 99.1% and 99.3% accordingly. The recall on the training set is

---

[85]Rammstein: Ich will
[86]Namika: Je ne parle pas français
[87]Adel Tawil: Bis hier und noch weiter

higher than on the test set — it amounts to 82.8%, which leads to 7.6% difference with the recall on the test set. The overall recall equals 76.6%.

Set	Precision	Recall
Training set	100%	82.8%
Test set	99.1%	75.2%
Total	99.3%	76.6%

Table 5.10: Precision and recall for extracting constructions of type *Verbs with prefixes*

Although there were only four false positives in a combination of the training and test sets, they could be split into three different categories of parsing errors produced by ParZu: an adverb analyzed as a participle (*bestimmt* in example 51a), a past participle analyzed as a finite verb (*gehört* in 51b is incorrectly analyzed as a finite verb of *gehören*) and a verb prefix analyzed as a past participle (*verrückt* in 51c is a prefix of the verb *verrücktspielen*). The first type of false positives occurs two times in the test set.

(51)   a. Das ganze Haus voller Frauen, ey, die schlafen **bestimmt**[88]

        b. Ich hab seit zwei Wochen nix mehr von ihm **gehört** [89]

        c. Bin ich es selber oder spielt die Welt **verrückt**?[90]

5.4.4 Passive voice

As for the previous topics connected to verb extraction, the training set for verb forms in the passive voice was also analyzed by the three parsers in order to re-evaluate their performance. The results of the analysis could be seen in table 5.11. Both ParZu and the Stanford parser showed the best result by analyzing 66.7% of the verb forms correctly, while the Mate tools succeeded in analyzing only twice as many verbs.

The distribution of errors made by each parser is shown in table 5.12. The following features had to be assigned correctly to extract constructions and generate distractors for the topic *Passive voice*: a part of speech, an infinitive of the verb, morphological features and a dependency relation. There are examples of errors involving all of these features.

---

[88]Cro: Whatever
[89]Lea: Zu dir
[90]Clueso: Neuanfang

Parser	Analyzed right
ParZu	**66.7%**
Stanford parser	**66.7%**
Mate tools	33.3%

Table 5.11: Percentage of the verb forms in the training set analysed correctly by each parser. Topic *Passive voice*

Parser	POS	Infinitive	Morph n.d.	Dep. rel.
ParZu	**16.7%**	0%	**16.7%**	0%
Stanford	**16.7%**	**16.7%**	0%	0%
Mate tools	0%	**50%**	0%	16.7%

Table 5.12: Distribution of errors for the analysis of the training set by each parser. Topic *Passive voice*

ParZu experienced difficulties both with determining a part of speech and assigning morphological characteristics. For instance, it did not assign a person to *werden* in 52a. For the Stanford parser, assigning the right infinitive and the right part of speech was equally complicated e.g. the parser analyzed the verb *gelangweilt* in example 52b as an adjective. The greatest number of errors by the Mate tools was connected to determining the infinitive of a verb. For example, *interviewt* in 52c was assigned an infinitive *interviewt*, while the right option would be *interviewen*. Apart from that, the Mate tools was the only parser that made an error in specifying a dependency relation between an auxiliary verb and a past participle, such as *bin* and *perforiert* in example 52d. Instead of assigning tag *aux* (auxiliary) to the auxiliary verb, the parser tagged it with *pred* (predicate).

(52)   a. Badeanstalt - **werden** rausgeschmissen, denn wir werfen mit Matsch und zeigen Penis den Mädchen[91]

        b. Hallo, Hallo, bist du auch so **gelangweilt**[92]

        c. Ein Speer in meinem Unterleib, ich werde **interviewt**[93]

        d. Ich **bin** völlig **perforiert** von Amors Pfeilen[94]

Both ParZu and the Stanford parser showed the same results for the training set, and ParZu was chosen as a parser for extracting constructions of type *Passive*

---

[91]Zugezogen Maskulin: Plattenbau
[92]Revolverheld: Lass uns gehen
[93]Alligatoah: Narben
[94]Farin Urlaub: Herz? Verloren

```
[
  "VAFIN\t?|?|Past|Ind\t\t",
  "VVPP\t\t0\taux"
],
```

Figure 5.12: Example of a rule for extracting verbs in passive voice

voice. This choice can be explained by a few reasons. As has been mentioned in section 5.2.1, the Stanford parser has a special dependency relation between an auxiliary verb and a past participle in a construction in passive voice, which allows to differentiate a passive construction with an auxiliary *sein* from a verb form in the perfect tense. However, checking if the infinitive of a past participle is contained in the list of the infinitives that can be used with *sein* in the perfect tense[95] would also help to detect constructions in the passive voice. The list already showed rather good results at differentiating verb forms in the perfect tense and verb forms in the passive voice in subsection 5.4.2. Therefore, together with the analysis of ParZu, the list could be used for extracting verb forms for the current topic. Moreover, choosing ParZu would allow to reuse some code written for extracting constructions for other topics, such as class *Word*.

All the rules for extracting verb forms in the passive voice looked very similar to the rules used for retrieving constructions in the perfect and pluperfect tense for the topic *Verb conjugation* — for instance, the rule shown in figure 5.12. However, in contrast to the blacklist for the topic *Verb conjugation* that included *werden* to avoid extracting verb forms in passive voice, the blacklist for *Passive voice* included *haben*, in order to ignore constructions in the perfect or pluperfect tense. The blacklist for surrounding words helped avoiding constructions with a spoken form of an auxiliary verb, such as *werd' erkannt* in example 53. The set of rules did not vary from the training to the test set.

(53)   Denn ich **werd'** nie **erkannt**, hier wurde Liebe verbannt[96]

The extraction of verb forms in the passive voice from the training set led to 3 false negatives and no false positives. The results on the test set included 8 false negatives and only 1 false positive which is shown in example 54. Instead of being analyzed as an infinitive, the verb *vergessen* is assigned a tag *VVPP* (past participle)

---

[95]See subsection 5.4.2 for the details on compiling the list.
[96]Maeckes: Niemandsland

and an infinitive *veressen*. Thus, the verb form in the future tense was understood as a construction in the passive voice.

(54)   Wir **werden** Euch niemals **vergessen**[97]

Set	Precision	Recall
Training set	100%	66.7%
Test set	98.6%	89.6%
Total	98.7%	87.2%

Table 5.13: Precision and recall for extracting constructions of type *Passive voice*

The results are given in table 5.13. The precision equals 100% for the training set, being also rather high for the test set. The precision for the combination of the training and test set amounts to 98.7%. The recall, on the other hand, is higher for the test set, leading to the difference of almost 23% between the training and the test set. The total recall equals 87.2%.

---

[97]Unheilig: Zeit zu gehen

6 Generating Distractors

6.1 Prepositions and their cases

The main goal of multiple-choice questions for topic *Prepositions and their cases* is, of course, helping German L2 learners to memorize the cases which are required by prepositions. Thus, in all types of prepositional phrases, words that illustrated these cases had to be substituted with blanks. To provide the user with a number of options for filling the blanks, distractors were generated.

Most words that were substituted with blanks in topic *Prepositions and their cases* belonged to closed word classes[1]. For such words, possible distractors did not need to be generated every time and could be manually listed. Closed word classes included possessive pronouns, personal pronouns, articles, demonstrative pronouns, indefinite pronouns and contractions of prepositions with definite articles.

Possessive pronouns have a separate form for every case and number. In singular number, a form also depends on the gender. However, every pronouns has only six unique surface forms. For instance, for a possessive pronoun *sein* these forms include *sein*, *seine*, *seinen*, *seinem*, *seiner* and *seines*. Therefore, a form of any possessive pronoun — *mein*, *dein*, *sein*[2], *ihr*[3], *Ihr*, *unser* and *euer* — always has five distractors, as the right option is excluded from the list of possible distractors. For instance, *deinem* in example 55 has the following distractors: *dein*, *deine*, *deinen*, *deiner*, *deines*.

(55) [Doch dann küsst du mich] Und ich tanz, in **deinem** Licht[4]

Personal pronouns get assigned two types of STTS-tags: *PRF*, which stands for a reflexive personal pronoun (e.g. *dich* in 56a) and *PPER* meaning a non-reflexive personal pronoun (e.g. *uns* in 56b). This differentiation, however, was not important at the stage of generating distractors.

---

[1]`https://en.wikipedia.org/wiki/Part_of_speech#Open_and_closed_classes`
[2]Corresponds both to *er* and *es*.
[3]Corresponds both to *sie* of singular number and *sie* of plural number.
[4]Andrea Berg: Diese Nacht ist jede Sünde wert

(56)   a. Hab' ich kurz nicht an **dich** gedacht[5]

       b. Kann mir irgendwer sagen, ob das mit **uns** klargeht?[6]

As no preposition can require a form of nominative case, only forms of accusative and dative case were included into lists of distractors for personal pronouns. Forms of pronouns *ihr* and *wir*, therefore, could not have any distractors, as their forms of accusative and dative cases are homonymous. For instance, *euch* in example 57a is a form of accusative case, and in example 57b — a form of dative case. For *sie*, distractors include forms of accusative and dative cases of both singular and plural number: *sie*, *ihr*, *ihnen*.

(57)   a. [Für die Familie und Freunde, seid nicht enttäuscht] Denn eigentlich leb' ich seit Jahren nur für **euch**[7]

       b. Ich will, dass keiner von **euch** sich ein Beispiel nimmt daran[8]

Thus, every personal pronoun, except for the excluded *ihr* and *wir*, has one distractor, as the right option is deleted from the list of distractors. For example, *dich* in example 56a has a distractor *dir*.

As possessive pronouns, both definite and indefinite articles in German have six unique forms — for instance, for the definite article, these forms include *der*, *die*, *den*, *dem*, *das* and *des*. Therefore, *die* in example 58 was assigned the following distractors: *der*, *den*, *dem*, *das* and *des*.

(58)   [Wolltest ihr was spendieren aber hattest dann nur den Fuffie] Der reichen sollte für **die** nächsten 20 Tage[9]

The members of the next two classes depended on lemmas found in the corpus of song texts. The only demonstrative found in the texts was *dieser*, so it was not necessary to include distractors for numerous other demonstratives, such as *jener* or *derjenige*. The STTS tagset distinguishes attributive indefinite pronouns and indefinite pronouns with determiners. Attributive indefinite pronouns used in the song texts include *irgendeiner* and *keiner*, while *viele* and *alle* are instances of indefinite pronouns with determiners found in lyrics. Surprisingly, both classes include the indefinite pronoun *jeder*. Different tags were assigned in very similar

---

[5]Von Wegen Lisbeth: Bitch

[6]Lea: Immer wenn wir uns sehn

[7]Lance Butters: 30

[8]Trailerpark: Bleib in der Schule

[9]Frittenbude: Bilder mit Katze

contexts — for instance, *jedem* in example 59a was assigned tag *PIAT*, corresponding to an attributive indefinite pronoun, and *jeder* in example 59b was analyzed as an indefinite pronoun with a determiner.

(59)   a.   Du stehst auf mit **jedem** neuen Tag

         b.   Hinter **jeder** neuen Tür [Kann die Sonne wieder schein'n][10]

Unique forms of the demonstrative *dieser* include *dieser*, *diese*, *dieses*, *diesem*, and *diesen*. Attributive indefinite pronouns *irgendeiner* and indefinite pronouns with a determiner *alle* and *viele*, as well as the indefinite pronoun *jeder* follow a similar declination pattern. The attributive pronoun *kein*, however, declines as possessive pronouns and has the following unique forms: *kein*, *keiner*, *keines*, *keinem*, *keinen*, *keine*. Thus, *diesen* in example 60a was assigned the following distractors: *dieser*, *diese*, *dieses*, and *diesem*. For *viele* in example 60b, distractors included *vieler*, *vieles*, *vielem*, and *vielen*. Words that the user could choose instead of *kein* in example 60c included *keiner*, *keines*, *keinem*, *keinen*, *keine*.

(60)   a.   [Ich will] Nach all **diesen** Jahren immer ohne Plan fahren[11]

         b.   Ich ließ über so **viele** Dinge Gras wachsen, dass ich's langsam wieder mähen muss[12]

         c.   Ich mache Rap denn ich hab Herz, keine Zeit für **kein** Gewimmer[13]

As we target acquiring standard German language in the app, it was decided to exclude non-standard contractions of prepositions with definite articles, such as *auf'm* in example 61. For this reason, no distractors could be found for some standard contractions — for example, the only possible distractors for *aufs* are *auf'm* and *auf'n*.

(61)   Meine 100 Enkel spielen Cricket **auf'm** Rasen[14]

Most contractions of prepositions with definite articles have one distractor — for instance, the distractor for *vorm* in example 62a is *vors*. A few contractions have two distractors. For instance, distractors for *unterm* in example 62b include *unters* and *untern*.

---

[10]Both examples are taken from the song *Ist da jemand* by Adel Tawil
[11]Clueso: Achterbahn
[12]Adel Tawil: Bis hier und noch weiter
[13]Haze: In der Luft
[14]Peter Fox: Haus am See

(62) a. [Ein Hotdog unten am Hafen] Und **vorm** Einschlafen schnell noch ein Bier[15]

  b. [Und wie jedes Jahr am Silvesterabend] Trinken wir auf unser Leben **unterm** Tellerrand[16]

As adjective is an open word class, all possible options could not be listed as easily as for other parts of speech. Therefore, another method for generating distractors was used. Although stems of adjectives differ from adjective to adjective, the endings stay the same in the same forms — for instance, *guter* in *guter Rat* and *neuer* in *neuer Tag* share the ending *-er*, as both adjectives have the form of singular number, masculine gender and nominative case. The unique adjective endings include *-er*, *-es*, *-en*, *-em*, *-e*. After the ending of the adjective used as a right option was identified, the rest of the adjective without the ending was used for generating distractors by appending all the possible endings to it. The ending of the right option had been excluded from the list of possible endings. In this way, for instance, distractors for *müdem* in example 63 were generated. They included *müder*, *müdes*, *müden*, and *müde*.

(63) [Während andre Hände schütteln aber hinterm Rücken haten] Bringe ich mit **müdem** Lächeln endlich Frieden für die Szene[17]

Since for every fill-in-the-blank question there should be either one or three distractors, there was sometimes the need to choose only a few distractors from the list of possible options. Although learning cases required by prepositions is important, learning gender of nouns also plays a great role in acquiring German language. For this reason, the number of distractors in prepositional phrases with nouns was limited by implementing a random choice.

6.2 Verb conjugation

In topic *Verb conjugation*, sometimes rather different strategies were used to generate distractors for verbs of various tenses and modes. For this reason, this section is split into several parts focusing on distractors for different verb forms.

---

[15]Element of Crime: Ein Hotdog unten am Hafen
[16]Adam Angst: Splitter von Granaten
[17]Audio88: Gnade

6.2.1 Perfect tense

Although it was clear which word had to be substituted with a blank when a verb form consisted of only one word, in case of verb forms in perfect tense, when a form was made of an auxiliary verb and a past participle, the following question arose: which part of the verb form should serve as a blank?

To answer that question, texts of German L2 learners from the MERLIN corpus[18] were analyzed. The corpus provides an access to texts written by learners of several European languages, including German. When an auxiliary verb preceded a past participle, the contexts where the distance between the auxiliary verb and the past participle varied between zero and five words were taken into account. If a perfect verb form started with a past participle, only contexts where there were no words between the two parts of the verb form were analyzed. In contexts with a different number of words separating two parts of a verb form in the perfect tense, the past participle was often not a dependent of the auxiliary verb anymore.

In total, 83 mistakes in perfect verb forms were found in the above described contexts. While 38 mistakes were connected to using the right form of an auxiliary verb, forms of past participles caused difficulties for learners in 45 cases. Therefore, generating distractors for past participles would bring more benefits to users of the app than creating distractors for auxiliary verbs.

Different types of mistakes in past participles were observed. One of the common mistakes was connected to using double letters in past participles. For instance, in example 64a, the author of the text used only one *m* instead of using two, as in the right past participle from verb *kommen* — *geko**mm**en*. In nine out of ten words with mistakes of this type, one of the double letters was missed. Only in one of the contexts, a double letter was written in a word which should be written with a single letter.

Another type of mistakes was incorrect usage of umlauts in past participles. In two out of three mistakes of this type, writing no umlaut when one was needed was observed, as in example 64b, where *gef**u**hrt* was used instead of *gef**ü**hrt*. In one of the contexts, however, a redundant umlaut was used.

Apart from these mistakes, three mistakes where an irregular verb form was built from a regular verb were found. For instance, an irregular verb form *gelach**en*** was used instead of the right form *gelach**t*** in example 64c. In four other cases, a regular verb form was used instead of an irregular one. Example 64d illustrates this type of

---

mistake with verb form *gebleibt*, used instead of *geblieben*.

(64)   a. Mein Eltern sind zu uns **gekomen**.[19]

       b. . . . Präsident der Informatik-Club, und habt viele Projecte **gefuhrt**.

       c. Wir haben viel getanzt, **gelachen** und gegessen und . . .

       d. Ich war krank und hat immer Tag ins Bett **gebleibt**.

       e. . . . und ich bin lange Zeit allein zu Hause **gebleiben**.

       f. Mein Mann hat endlich einen Arbeitplatz **gefundet** und ich habe nach zwei . . .

There are two more types of mistakes associated with irregular past participles with a vowel change. One of the types involves using an irregular participle without a vowel change, such as *gebleiben* used instead of *geblieben* in example 64e. In the other type, the vowel change is preserved, but the ending of a regular participle is used e.g. *gefundet* in example 64f. The right form has the ending of an irregular past participle *-en*. There are in total four mistakes of these two types.

Twenty-three other mistakes did not display enough similarities to be grouped together and were classified as miscellaneous. All the types of mistakes described above were used for creating distractors for past participles.

The types of distractors generated for a past participle to a certain extent depended on the group that the verb belonged to. The classification from the "Hammer's German grammar and usage" was used (Durrell, 2017). The author divides verbs into regular weak verbs (e.g. *kaufen*), irregular weak verbs (e.g. *kennen*) and strong verbs, all of which are irregular (e.g. *gehen*). To determine which group a verb belongs to, it was first checked if its past participle ended with *-t*. If it had another ending, the verb was considered strong as e.g. *gehen* with the past participle *gegangen*. Otherwise, the verb was either a regular or an irregular weak verb. To distinguish these two groups, the part of the infinitive of the verb without the ending *-en* or *-n* and the part of the past participle without the ending *-et* or *-t* were compared. Before the comparison, the prefix *ge-* had to be deleted from the past participle, if it contained one. In this way, verbs *kaufen* and *kennen* with the past participles *gekauft* and *gekannt* are assigned labels *regular weak verb* and *irregular weak verb* accordingly, as the part of the infinitive *kauf-* coincides with the part of the past participle, and the part of the infinitive *kenn-* is different from the part of the past participle *kann-*.

---

[19]All examples in 64 are taken from the MERLIN corpus. The original spelling is preserved.

For regular weak verbs, an irregular participle without a vowel change and an irregular participle with a vowel change were generated as distractors. The second type of distractors is the only type not supported by mistakes from the MERLIN corpus. Using an irregular participle without a vowel change instead of a regular form is illustrated in example 64c. To generate such distractor, the ending of a regular participle *-t* or *-et* was replaced with the irregular ending *-en*. For instance, from a regular participle *abgefeiert* in example 65, *abgefeier***n** was generated. A necessary condition for generating a distractor of this type was prefix *ge-* in the regular participle of a verb. Without the prefix, the form of the distractor would be the same as the infinitive of the verb and would not serve as a decent distractor e.g. *verkauf***en** created from *verkauft*.

(65) Hauptsache Audiolith, habt alles **abgefeiert**[20]

An irregular participle with a vowel change was generated with patterns of irregular verbs. To create the patterns, all the irregular verbs listed in the "Hammer's German grammar and usage" were written out (Durrell, 2017). The data provided by Morphy, a tool for morphological analysis of German texts[21], was used to complete the list of irregular verbs. The data from Morphy contains lines describing word forms of German words. Every line consists of a word form, a lemma of the word and tags denoting the part of speech and grammatical characteristics of the word form. Irregular past participles were identified by tag *VER:PA2:NON*.

The irregular verbs in the list were grouped together based on combinations of vowel changes and participle endings. For instance, *fliegen* (*flog, geflogen*) and *schieben* (*schob, geschoben*) belong to one group with vowel changes *ie - o - o* and an ending *-en* for a past participle. Seventeen groups with minimum of two verbs pro group were formed. Some verbs did not belong to any group e.g *bitten* (*bat, gebeten*). The groups were further split into verbs with similar endings. Thus, *fliegen* together with *biegen* and *wiegen* displayed a pattern *iegen - og - ogen*[22] and verbs *schieben* and *stieben* — *ieben - ob - oben*. Some patterns were applicable to only one irregular verb, such as *ingen - ang - ungen* for *zwingen*. In total, 68 patterns were formed.

In order to find a suitable pattern for a regular verb, all the prefixes should be cut from its infinitive first. After that, it is checked if any pattern of infinitive, such as *iegen* or *ieben*, with one or more consonants before it matches the whole infinitive

---

without prefixes. If such pattern has been found, an irregular past participle could be formed by attaching the ending corresponding to the past participle for this pattern to a part of the regular participle. For example, in order to generate a distractor for a regular participle *erlebt* in example 66, a pattern suiting its infinitive without a prefix — *leben* — had to be found. A pattern *eben - ob - oben* seemed to be a good candidate. To attach the ending of an irregular past participle *oben*, the regular ending *-t* and a part of the infinitive pattern *eb* was cut from the regular participle *erlebt*. In this way, an irregular distractor *erloben* was derived.

(66)　Neulich hat ein Kumpel was **erlebt**[23]

Irregular weak verbs could have three types of distractors: a regular participle without a vowel change, an irregular participle without a vowel change and an irregular participle with a vowel change. Using a regular participle without a vowel change instead of the right irregular form is illustrated by example 64d, repeated here in example 67.

(67)　Ich war krank und hat immer Tag ins Bett **gebleibt**.

Although the examples found in the corpus include such mistakes only for strong verbs, we believe that the strategy could also be applied to generating misleading distractors for weak irregular verbs. To generate such distractor, the ending *-en* or *-n* was cut from the infinitive of the verb first. Then, if the verb was splittable and the irregular participle contained a prefix *ge-*, this prefix was inserted before the root of the infinitive without the ending. If the stem ended with *-t* or *-d* after any letter, or *-n* or *-m* after a letter denoting another consonant, the letter *e* had to to be added to the stem before finally adding the ending of a regular past participle *-t*. For instance, to generate a regular distractor with a vowel change for *gebracht* in example 68, *-en* was cut from the infinitive of the verb *bringen*. Then, as the past participle *gebracht* contains prefix *ge-*, it was attached to the part of the verb's infinitive to get *gebring*. The final regular participle without a vowel change was *gebringt*.

(68)　Es hat leider nichts **gebracht**[24]

The process of generating an irregular participle without a vowel change has already been described for regular weak verbs. To generate such distractor for

---

[23]Adam Angst: Professoren
[24]Farin Urlaub: IDisco

irregular weak verbs, their distractor of type *regular participle without vowel change* was used. Thus, in case of *gebracht* from example 68, *gebringt* is used to generate an irregular participle *gebringen*.

The generating of an irregular participle with a vowel change by using patterns of irregular verbs has also been already explained before. The infinitive of past participle *gebracht* in example 68 corresponds to an irregular pattern *ingen - ang - ungen*. Therefore, a new distractor *gebrungen* was derived.

Distractors for strong verbs include three already mentioned types of distractors: a regular participle without a vowel change, an irregular participle without a vowel change, and an irregular participle with a vowel change. Moreover, there is one type of distractors generated exclusively for strong verbs — namely, a regular participle with a vowel change. Using such participle instead of the right form of a strong verb could be observed in example 64f, repeated here in example 69.

(69) Mein Mann hat endlich einen Arbeitplatz **gefundet** und ich habe nach zwei . . .

To generate a regular participle with a vowel change, the ending *-n* or *-en* was cut from the irregular past participle. After that, the vowel *e* was attached to the stem, if needed, and the ending of a regular past participle *-t* was added. In this way, a regular participle with a vowel change *angezogt* could be created for *angezogen* in example 70.

(70) Du hast mich **angezogen**, ausgezogen, großgezogen[25]

Examples of other types of distractors for a past participle of a strong verb include *angezieht* (a regular participle without a vowel change), *angeziehen* (an irregular participle without a vowel change) and *angezohen* (an irregular participle with a vowel change), generated using the pattern *iehen - oh - ohen*. If the form coinciding with the right past participle was generated using the patterns of irregular verbs, it was deleted from the final list of distractors.

Apart from distractors depending on the type of a verb, there were two types of distractors independent from it. One of the strategies involved replacing double letters denoting consonants with a single letter, following the mistake in 64a, repeated here in 71. The reverse case was only found in one out of ten examples from the corpus and was, therefore, not implemented.

---

[25]AnnenMayKantereit: Oft gefragt

(71)   Mein Eltern sind zu uns **gekomen**.

Some limitations were set on generating a distractor containing this type of mistake. Firstly, if double letters appeared several times in a past participle, only one such combination was replaced with a single letter. Secondly, the double letters that got replaced had to be located in the root of the verb and not in its prefix. However, the first part of some compound verbs could also contain a double letter e.g. *bekanntgegeben*. Such cases could not be avoided, as well as cases with verbs with prefixes like *misslingen*[26]. Distractors where double letters were replaced with a single letter were created using all distractors generated so far, as well as the right past participle. For instance, for a regular past participle *gestellt* in example 72 distractors *gestollen* and *gestellen* were generated. Therefore, after replacing the double letters, the following distractors were created: *gestelt*, *gestolen* and *gestelen*.

(72)   Du hast die Tastentöne laut **gestellt**[27]

Another strategy of generating distractors for all types of verbs consisted of replacing umlauts with the corresponding letters denoting vowels without sound change and replacing letters denoting vowels without sound change with the corresponding umlauts, if there were any. To put it simply, as there are only three umlauts in the German language, *o*, *a* and *u* got replaced with *ö*, *ä* and *ü* correspondingly, and *ö*, *ä* and *ü* were changed into *o*, *a* and *u*. A context from the corpus illustrating this mistake was given in example 64b, repeated here in example 73.

(73)   . . . Präsident der Informatik-Club, und habt viele Projecte **gefuhrt**.

After analyzing the mistakes in the corpus and the results received during the development of the strategy, a number of limitations was introduced. As with double letters, a letter was replaced only if was located in the root of a verb. However, as in case of double letters, the first part of a compound verb also sometimes went under replacement, resulting in distractors like *bekänntgegeben*. Apart from that, there had to be only one letter that could be replaced. Moreover, the letter to be replaced had to be surrounded by letters denoting consonants from both sides, so that diphthongs like e.g *eu* in *gestreunt* could be avoided. The replacements were only allowed in a regular participle without a vowel change or

---

[26]More on verb prefixes and ways to identify prefixes in a verb in section 6.3.
[27]Audio88: Halleluija

in an irregular participle without a vowel change. For instance, for a past participle *betrogen* in example 74a only distractor *betrügt* from its distractor in form of regular past participle *betrugt* was generated, as we believe that replacing letters denoting vowels in past participles with a vowel change makes the resulting distractors highly improbable for a student e.g. *betrögen* or *betrögt*. For *getragen* in example 74b, two distractors with umlauts were generated: *geträgt* and *geträgen*.

(74)   a. [Ich war noch nie so weit am Boden Denn ich weiß, mein ganzes Geld ist weg] Und meine Frau hat mich **betrogen**[28]

          b. Ihr habt uns mit Applaus **getragen**[29]

The final list of distractors did not contain any past participles that exist as real forms, as that could mislead the student. For instance, from distractors for *gewunken* in example 75 — *gewunkt*, *gewinken* and *gewinkt* — *gewinkt* was deleted, as it is one of the possible past participles for *winken*. The list of real past participles is obtained from the data provided by Morphy.

(75)   Früher warst du immer hier und hast **gewunken** bis zum Schluss[30]

6.2.2 Past tense

To assign distractors to verbs in the past tense, the verbs had to be divided into regular and irregular verbs. The group of irregular verbs includes both weak irregular verbs and strong verbs. To distinguish regular and irregular verbs, it was checked if the infinitive of the verb without the ending *-en* or *-n* served as the beginning of the past form. For instance, *kaufen* belongs to regular verbs, as the past form *kaufte* starts with the part of the verb's infinitive *kauf*. The past forms of irregular verbs *kennen* and *gehen*, however, are both different from the root of their infinitives: *kannte* does not start with *kenn* and *ging* — with *geh*. Although the examples are given for the verb forms of singular number and 1ˢᵗ person, the method works for other past forms, too.

Past forms of regular verbs were assigned irregular distractors. Similarly to irregular past participles, such irregular past forms were generated using patterns of irregular verbs. For instance, in order to generate a distractor for *sagtest* in

---

[28]Cro: Whatever
[29]Unheilig: Zeit zu gehen
[30]Philipp Dittberner: Jede Nacht

example 76, a pattern *agen - ug - agen - ägst/ägt*[31] matching the verb was found. To get a past form of 1st person and singular number, the infinitive pattern had to be cut from the verb's infinitive, and a past pattern had to be appended instead. Thus, *agen* was cut from *sagen* to get *s-*, and after attaching *-ug* to it, a distractor *sug* was generated. However, getting other forms of the irregular past form is also useful, as it would provide us with more suitable distractors and mislead the user even more.

(76)   Du hast sie angesprochen, **sagtest** neckisch: Ey Katze![32]

In the past tense, the forms of 1st and 3rd person are the same both for singular and plural number. Therefore, any verb can have maximum four unique forms of past tense. From *sug*, the following forms were generated: *sugt, sugst, sugen*.

Distractors for an irregular verb included regular forms of past tense and irregular forms that were different from the real past forms of the verb. To generate a regular past form, the ending *-en* or *-n* was cut from the infinitive of the verb. After that, all the regular endings — *-te*, *-test*, *-ten*, *-tet* — were added to the part of the infinitive, in order to get the four unique regular forms of past tense. The following distractors were generated for *stand* in example 77 by attaching regular endings to *steh-*: *stehtest, stehten, stehte, stehtet*.

(77)   Wo mal dein Fahrrad **stand**, liegt jetzt nur dein Schloss im Graben[33]

In a similar way as for regular verbs, irregular past forms could be generated for irregular verbs. For instance, *stehen*, the infinitive of the verb *stand* in example 77, corresponds to the pattern *ehen - ah - ehen - ieht*. After the pattern was found, *ehen* was cut from the infinitive, and the combination of *st-* and *-ah* resulted in an irregular distractor *stah*. Other forms generated from the distractor include *stahen*, *stahst* and *staht*.

During the generation of both regular and irregular past forms, it can happen that the letter *e* should be attached to the *stem* — a part of an infinitive without the ending for regular forms or a past form of 1st person and singular number for irregular forms. If the stem ends with *-t* or *-d* after any letter, or *-n* or *-m* after a letter denoting another consonant, *e* is inserted between the stem and the ending. For instance, in order to generate regular distractors for *fanden* in example 78, *e* should

---

[31]Some irregular verbs have vowel changes in forms of 2nd and 3rd person, singular number, present tense e.g. *schlägst* and *schlägt* of *schlagen*. The forth part of the pattern contains endings given for generating such verb forms.

[32]Frittenbude: Bilder mit Katze

[33]Zugezogen Maskulin: Plattenbau

be inserted between *find-* and endings of regular past forms to get the following distractors: *findeten*, *findete*, *findetet*, *findetest*.

(78)   Wir **fanden** einen Schatz, inmitten von Schrott[34]

Apart from distractors that depended on a verb being regular or irregular, other past forms of a verb from the song texts were added to the list of its distractors. To make it possible, past forms were extracted from the data provided by Morphy. Every infinitive found in the data was assigned *regular* conjugation type, *irregular* conjugation type, or both, and past forms for the type were sorted based on their person and number. All the information was stored in a Python dictionary. Therefore, if the infinitive of the verb was present in the data, all its forms corresponding to the conjugation type of the verb could be extracted. After the past form of the verb used in the song text was deleted from the list of the past forms, the rest of the past forms could be used as distractors. For example, other past forms of *finden* — *fandet*, *fand*, *fandest* — were extracted from the dictionary of past forms to be used as distractors for the verb *fanden* in example 78.

As with distractors generated for past participles, it was checked that all the synthesized distractors for past forms did not coincide with past forms of actual verbs. To carry out this check, all the past forms found in the Morphy's data were collected. For instance, irregular distractors *log*, *logt*, *logen* and *logst* generated for *lag* in example 79 are deleted from the list of distractors, as they constitute real forms of the verb *lügen*.

(79)   Drunter **lag** ein Mann, der seine Hand nach uns streckt[35]

6.2.3 Present tense

Not all verbs that have irregular past participles and irregular past forms also have irregular forms of present tense e.g. an irregular form *hilft* of *helfen* and a regular form *fliegt* of *fliegen*. Moreover, for most irregular verbs, only present forms of 2nd and 3rd person and singular number are irregular. For instance, although *helfen* has the irregular present forms *hilfst* and *hilft*, other present forms of the verb — *helfe*, *helfen* and *helft* — are regular. Therefore, it is impossible to determine if a verb has a regular or an irregular conjugation in the present tense when a form of 1st person singular number or any form of plural number is encountered. For

---

[34]Frittenbude: Einfach nicht leicht
[35]Alligatoah: Lass liegen

this reason, every verb gets assigned both irregular and regular distractors that get subsequently filtered out.

As for past participles and past forms, irregular distractors for present forms were generated with the patterns of irregular verbs. However, as has been mentioned above, only some irregular verbs have irregular forms of present tense. Out of 68 irregular patterns, there are 32 such patterns. Some patterns have endings for forms of both 2$^{\text{nd}}$ and 3$^{\text{rd}}$ person (e.g. *ergen - arg - orgen - **irgst/irgt***) and for some, both forms have the same ending (e.g. *ersten - arst - orsten- **irst***). To generate irregular distractors, a suitable irregular pattern had to be found first. Then, if the pattern had endings of the present tense, the pattern corresponding to the infinitive was cut from the infinitive of a verb, and all the possible present endings were combined with the part of the infinitive. For instance, to generate irregular distractors for *schmeckt* in example 80, the pattern *ecken - ak - ocken - ickst/ickt* was used. Cutting *-ecken* from the infinitive of the verb resulted in *schm-*, and present forms *schmickt* and *schmickst* were generated by appending endings of present forms to it.

(80) Die Kippe **schmeckt** nach Liberté [Solang wir beide sie uns teil'n][36]

As a set of irregular present forms and a set of regular present forms only differ in forms of 2$^{\text{nd}}$ and 3$^{\text{rd}}$ person and singular number, maximum two regular distractor could be generated. The process of generating regular distractors started from cutting endings *-en* or *-n* from the infinitive of a verb. Then, if the part of the infinitive ended with *-t* or *-d* after any letter, or *-n* or *-m* after a letter denoting another consonant, letter *e* was added to the part of the infinitive. Finally, by appending endings of the present tense to it, one or two regular present forms were generated. If a stem of a real regular verb ends with *z, ß, x* or *s*, the present forms of the 2$^{\text{nd}}$ and 3$^{\text{rd}}$ person and singular number are homonymous. The same rule was applied to synthesizing regular forms — for example, *sprichst* in example 81a was assigned two regular distractors, *sprecht* and *sprechst*, and for *weiß* in example 81b, only *wisst* could be generated, as part of the infinitive without the ending *wiss-* ended with *s*.

(81) a. Auf einmal **sprichst** du mich an[37]

     b. Ich **weiß** schon, du heißt Isi, aber ist mir egal[38]

---

[36]Namika: Je ne parle pas français
[37]Namika: Je ne parle pas français
[38]Cro: Easy

76

After regular and irregular distractors were generated for a verb form in present tense, all real present forms were deleted from the list of distractors. Therefore, actual present forms of the verb and real present forms of other verbs were not present in distractors anymore.

As the next step, other present forms of the verb were added to the list of distractors. For every verb, the search was conducted first for an irregular, and then for a regular conjugation type. The search for the regular conjugation type was only conducted if no forms were extracted for the irregular type. Present forms for a conjugation type were only extracted if the verb form from the song texts was present in the list of forms for this conjugation type. For instance, if distractors for a verb *milkt* were needed, the forms belonging to the irregular conjugation type would be extracted, such as *melken*, *melke*, *melkt* and **milkst**. For a verb form *melkt*, on the other hand, regular present forms would be extracted: *melken*, *melke* and **melkst**. If a verb form belonged both to regular and irregular present forms, irregular present forms were extracted, as the search for the irregular conjugation type was conducted first.

For creating both the list of present verb forms for deleting real present forms from the list of distractors and a Python dictionary for extracting other present forms of verbs, the data provided by Morphy was used.

Sometimes the infinitive of a verb was not found in the Python dictionary containing present forms of verbs, so no other present forms of the verb could serve as distractors. If, in this case, the form of the verb did not coincide with the infinitive, the infinitive was added to the list of distractors to increase the number of distractors. As an infinitive usually coincides with the present forms of 1st and 3rd person and plural number, it can serve as a suitable distractor. For example, for *cruise* in the line presented in 82, only one regular distractor *cruist* could be generated, since there was no irregular pattern suitable for the verb, and no present forms corresponding to *cruisen* could be found in the Python dictionary. Adding *cruisen* to the list of distractors made the choice for the student to a certain extent more difficult.

(82)   Ich **cruise** mit meim' Cousin in nem Cadillac umher[39]

---

[39]Haze: In der Luft

6.2.4 Imperative mode

When an imperative form of plural number was encountered, it was impossible to determine if a verb has a regular or an irregular conjugation type in the imperative mode. For instance, *helft* does not signal that verb *helfen* has an irregular imperative form of singular number. Therefore, as for verbs in the present tense, verbs in the imperative mode were assigned both regular and irregular distractors that were filtered out afterwards.

As one could infer from the previous paragraph, irregular forms in the imperative mode exist only for singular number. Thus, only such forms were generated as irregular distractors. As irregular distractors in the indicative mode, irregular distractors in the imperative mode were also created using patterns of irregular verbs. However, there are even less irregular patterns that are assigned irregular forms in the imperative mode than there are patterns exhibiting irregularities in the present tense. Only 17 out of 68 patterns could be used for generating irregular forms in the imperative mode. If one of these 17 patterns suited the verb form from the song texts, the pattern of the infinitive was cut from the infinitive of the verb and the ending of the imperative form of singular number was attached to the rest of the infinitive. For instance, a suitable pattern was found for verb *geh* in example 83 — *ehen - ah - ehen - ieht - ieh*[40]. An irregular distractor *gieh* was generated by cutting *ehen* from *gehen* and attaching the last ending of the pattern — *-ieh* — to *g-*.

(83)   Komm, nimm meine Hand und **geh** mit mir![41]

As with irregular distractors, it was reasonable to create only one form of regular distractors. Regular distractors were generated differently for verbs with an infinitive ending with *-en* (e.g. *nehmen*) and verbs with an infinitive ending with *-n* (e.g. *liefern*). If the infinitive of a verb ended with *-en*, cutting the last letter from the infinitive created a regular form, such as *nehme* from *nehmen*. For verbs with the infinitive ending in *-n*, not only the last *-n* was cut from the infinitive, but also the letter *e* was added to it to acquire a regular distractor e.g. *leifere* for *liefern*. A synthesized form *gebe* is an example of a regular distractor for a verb from the song texts *gib*. The verb in context is shown in example 84.

(84)   **Gib** mir eine neue Idee[42]

---

[40]The last ending stands for the ending of imperative mode and singular number.
[41]Helene Fischer: Atemlos durch die Nacht
[42]Peterlicht: Neue Idee

For most German verbs, the ending *-e* is optional in the imperative form of singular number e.g. for *kommen* in example 85a and *setzen* in example 85b (Durrell, 2017). Although we create all the regular distractors with this ending, dropping it is usually not considered a mistake. For this reason, regular forms of imperative mode and singular number without the ending were extracted from the text. However, it was checked that for this kind of forms, no regular distractor was generated. For instance, for *lass* in example 85c, a regular distractor *lasse* was not used, as the verb form itself was a regular form of the imperative mode.

(85)   a. Komm(e) in den Garten!

      b. Setz(e) dich![43]

      c. **Lass** uns hier raus[44]

It was also checked that both an irregular and a regular distractor do not coincide with any real forms of the imperative mode. The distractors that did not comply to this rule were excluded. The list of the imperative forms was extracted from the data provided by Morphy.

The imperative forms stored in the Morphy data were also used for extracting other imperative forms for a verb form from the song texts. As verbs in the present tense, verbs in the imperative mode were not classified into regular and irregular. Therefore, both conjugation types were tried. The regular conjugation type was searched for first. If the verb form or its version with *e* attached (for verbs like *lass* in example 85c) was not found in the list of verbs belonging to the regular conjugation type, the irregular conjugation type was chosen. Since a regular and an irregular conjugation of every lemma could be assigned only two verb forms each — of singular and of plural number, only one distractor could be found for a verb form from the lyrics. For a verb of singular number, like *spar* in example 86a, its form of plural number was used as a distractor e.g. *spart*, and for a verb of plural number, a verb of singular number was extracted cf. *fürchtet* in example 86b and its distractor *fürchte*.

(86)   a. [Verkneif dir jegliches Mitleid] Und **spar** dir jedes Klischee[45]

      b. **Fürchtet** euch, fürchtet euch nicht[46]

---

[43]Examples 85a and 85b are taken from "Hammer's German grammar and usage" (Durrell, 2017)
[44]Revolverheld: Lass uns gehen
[45]Farin Urlaub: OK
[46]Rammstein: Sonne

6.3 Verbs with prefixes

For generating distractors for topic *Verbs with prefixes*, verbs sharing the same root had to be grouped. For creating such groups, a list of German prefixes and a list of infinitives were needed. A rather thorough list of German prefixes could be found in Wiktionary[47]. It contains the 320 prefixes, some of which are used exclusively for forming nouns e.g. *Wohn-* (*Wohnnort*, *Wohnsitz*). Such prefixes were not needed for working with verbs and were, therefore, excluded from the list.

For compiling a list of infinitives, two sources were used. One of the sources is the database of German verbs created as part of the thesis project by Sabine Buchholz, which has already been mentioned in section 5 (Buchholz, 1996). Another source is the data provided by Morphy. Both of these sources have a rather impressive number of infinitives. Moreover, one source tends to complete the other.

While the database of German verbs provides exclusively infinitives of verbs, the data made available by Morphy contains various word forms of different parts of speech. Infinitives were identified by tags *VER:INF:SFT* (for regular verbs) or *VER:INF:NON* (for irregular verbs). There were 11,998 unique infinitives found in the verb database, and 8,423 infinitives — in the data from Morphy. Since many of the infinitives from the verb database and the data provided by Morphy were the same, the total number of infinitives in the list compiled from the two sources was 13,381.

Using German prefixes extracted from Wiktionary, the infinitives from the list were divided into two groups: verbs that do not have any of the prefixes listed in Wiktionary and verbs that did not fall into the first category. After that, verbs with a single prefix that share the root with one of the verbs without any prefixes were detected in the second group. Two data structures were created to ease the process of finding verbs with prefixes that share the same root. One of the dictionaries contains verbs with a single prefix as keys and the corresponding verbs without prefixes as values. For instance, *nachreifen* could be a key, and *reifen* — its value. The other dictionary consists of lists of verbs that have a prefix and share the same root and the corresponding verbs without any prefixes. In this dictionary, *reifen* could be a key with the following list of verbs being its value: *nachreifen*, *ausreifen*, *bereifen*, *heranreifen*, *umreifen*.

Thus, when a verb with a prefix was encountered in a song text, other verbs with prefixes that share the root with the verb were acquired in the following way. Firstly,

---

[47]https://en.wiktionary.org/wiki/Category:German_words_by_prefix

the verb without the prefix corresponding to the verb from the lyrics was found in the first dictionary. Secondly, using the verb without the prefix as a key in the second dictionary, a list of verbs sharing the same root with the verb was extracted. After the verb was deleted from that list, its members could be used as the basis for creating distractors.

The method of grouping verbs described above worked only for the verbs with prefixes that had equivalent verbs without the prefixes. Therefore, infinitives like *gelingen* and *misslingen* could not be grouped together, as it was impossible to state that they were related.

A verb with a prefix could have one of the following forms: a finite verb form, an infinitive, a verb in the imperative mode and a past participle. Distractors for the first three verb forms were generated by replacing the prefix of the right option with prefixes of other verbs that share the same root with the right option. For instance, example 87 contains a word *gefällt* of present tense, singular number and third person. The prefix *ge-* could be cut from the beginning of the verb form and replaced with e.g. *ent-*, *her-* and *miss-* to get distractors *entfällt*, *herfällt*, *missfällt*. The prefixes were extracted by comparing each verb with a prefix against the verb without prefixes.

(87)   [Und du drehst dich in deiner kleinen Welt] Mit deinen großen Farben, wie
        sie dir **gefällt**[48]

This way of generating distractors in the same form as the right option, however, was not suitable for past participles because of separable verbs. Unlike inseparable verbs, separable verbs require *ge-* to be inserted between the prefix and the root of the past participle. Some prefixes clearly signify that a verb is separable or inseparable — for instance, verbs containing prefix *auf-* are separable (*auf**ge**standen*) and verbs with prefix *be-* are inseparable (*bestanden*). Other prefixes can be both separable and inseparable, such as prefix *um-* in *um**ge**schaut* and *umarmt*. There are even cases when the same verb can be used as a separable or an inseparable verb e.g. ***um**fahren*. The meaning of verbs like *umfahren* depends on them being separable or inseparable.

Therefore, it was not possible to conclude if a verb was separable or inseparable from its infinitive, so it was not clear when the prefix *ge-* had to be inserted into past participles, and when not. It was decided to use the data provided by Morphy to find a past participle for every verb with a prefix contained in the two Python

_____

[48]Philipp Dittberner: In deiner kleinen Welt

dictionaries. In this way, when verbs had to take a form of past participle, their form could be found in the list of past participles extracted from Morphy. In case of verbs like *umfahren*, a past participle of either separable or inseparable verb could be used, as it did not affect its role as a distractor.

Example 88 illustrates the case when the verb with a prefix has a form of past participle. Using the two Python dictionaries, the infinitives *zusammenpassen*, *anpassen*, *abpassen* were collected. With the list of past participles, the following distractors were generated from them: *zusammengepasst*, *angepasst*, *abgepasst*.

(88)   Ich mach' alles das, was ich **verpasst** habe[49]

6.4 Passive voice

In order to decide which word of a verb form in the passive voice — the auxiliary verb or the past participle — had to be substituted with a blank, examples from the MERLIN corpus[50] were analyzed. The same contexts as for the verb forms in the perfect tense were selected. The authors of the texts in the corpus made 8 mistakes in auxiliary verbs and 8 mistakes in building forms of past participles. Therefore, the two parts of a verb form in the passive voice were considered equally hard for German L2 learners, and distractors for both auxiliary verbs and past participles were needed. For every verb form in the passive voice found in the song texts, it was randomly decided which of the two parts of the verb form would serve as a blank.

For generating distractors for past participles, the same strategy was used as for past participles that were part of verb forms in the perfect tense. For instance, for *gefroren* in example 89, the following distractors were generated: *gefriert*, *gefrort* and *gefrieren*. To choose types of distractors for auxiliary verbs, the mistakes from the corpus were analyzed.

(89)   Irgendwo sind sie **gefroren** [Unsere Spuren][51]

Two out of eight mistakes could not be classified. Two other mistakes were connected to using a form of *sein* instead of *werden*, as in example 90a, where *werden* should be written in place of *sind*. We supposed that mistakes where *werden* was used instead of *sein* were also possible.

---

[49]Mark Forster: Au Revoir
[50]https://merlin-platform.eu/index.php
[51]Frittenbude: Einfach nicht leicht

In the other four mistakes, the authors of the texts confused the number of an auxiliary verb. For instance, in example 90b, *wurden* is the right auxiliary.

(90)   a. Diese Personen **sind** als "Migranten" bezeichnet.[52]

        b. Die Schnellimbisse **wurde** an der Welt gegründet.

Based on the two types of mistakes, two types of distractors for auxiliary verbs were created. Two Python dictionaries were used for these purposes. In one of the dictionaries, to every form of *sein* in the present and the past tense, the corresponding form of *werden* was assigned, and to every form of *werden* — the form of *sein*. For example, form *werde* corresponded to *bin*, and *warst* was assigned to *wurdest*. The other dictionary contained forms of auxiliaries to which corresponded the same auxiliaries of the same tense, but of a different number. For instance, such key-value pairs as *bist* and *seid*, or *wurden* and *wurde* could be found in the dictionary. To verbs *sind* and *werden*, an array containing forms of 1st person and singular number and of 3rd person and singular number was assigned, as the verbs can correspond both to 1st and 3rd person. To get a list of distractors for an auxiliary verb, first the value from the second dictionary was extracted. Then, values from the first dictionary corresponding both to the extracted verb(s) and to the auxiliary verb from the song texts were added to the list of distractors.

(91)   Hier wird gebaut, **wird** noch geplant[53]

To *wird* in example 91, for instance, a distractor *werden* from the second dictionary was assigned. Then, using the first Python dictionary, verbs *sind* and *ist* were appended to the list of dictractors.

---

[52]All examples in 90 are taken from the MERLIN corpus. The original spelling is preserved.

[53]Bratze: Zitate

7 Developing the Application

7.1 Filtering extracted constructions

The extracted constructions of all topics had to be filtered out before being used for exercises in the app. A few principles governed this process. First of all, if one line contained several constructions that either began or ended with the same word, the longest construction was left, while the shorter construction was deleted. For instance, from the line in example 92, two prepositional phrases ending with the word *Gebäude* were extracted: *im gegenüber liegenden Gebäude* and *gegenüber liegenden Gebäude*. Since the second construction was shorter, it was filtered out.

(92)   Doch die Kripo sitzt **im gegenüber liegenden Gebäude**[1]

Another principle concerned the optimal number of constructions for the user to be able to answer the corresponding multiple-choice questions. It was believed that it would be too stressful for the user to answer several questions within one line of lyrics. If a line, therefore, contained more than one construction of the same type, only one of the constructions was left. The line in example 93 contained two finite verb forms: *kommt* and *nimmt*, with *nimmt* deleted from the list of constructions implemented in the exercises.

(93)   Der Teufel **kommt** und **nimmt** mich in den Arm[2]

To provide the user with more time to choose the right option for a multiple-choice question and make the process of performing an exercise more relaxing, it was decided that the learner should be allowed to answer the question as long as the line after the line with the blank is sung. This implied that if two lines placed next to each other contained the same type of constructions, the constructions extracted from the second line had to be filtered out. For instance, to give the user more time to fill the blank with a verb with a prefix *gehören* in the first line in example 94, *verzweifle* from the second line was not included in the exercise.

---

[1] Haze: Weisch Weisch
[2] Isolation Berlin: Alles grau

(94) Die **gehören** in die Wäsche, ist denn das so schwer? [Ich **verzweifle** an dir, ich glaub ich liebe dich nicht mehr][3]

Lastly, to make exercises more interesting for the student, we checked that every construction was found at most two times in a song text. If some constructions occurred several times, the two occurrences furthest from each other were left, while all other occurrences were deleted. For example, a passive construction *waren geboren* occurred ten times in the song text *Geboren um zu leben* by Unheilig. The lines where it was found had the following numbers: 11, 13, 25, 27, 31, 35, 37, 39, 41, 44. From these ten occurrences, the algorithm selected only the construction in line 11 and the construction in line 44.

To sum up, by applying different strategies for filtering out the constructions used in the exercises, we ensured that the learning process is not too stressful for the user and the multiple-choice questions that he/she encounters are diverse.

7.2 Database

A relational database was used to store the information needed for the operation of the app. As a relational database management system, SQLite[4] was chosen. There are libraries that enable work with SQLite databases both in Java and in Python. As the processing of song texts had been mainly carried out with Python, it was, therefore, very practical that the database could be created in Python. The app, on the other hand, was written in Java, and the ability to exploit the database with the library for Java also proved to be very useful.

The database contains two tables: a table with the information about songs and a table for storing constructions extracted from song texts. The table for songs includes a song id, the name of the song, the name of the artist performing the song, the genre the artist belongs to, the id of the song video on YouTube (the part selected in bold in example 95) and the thumbnail of the song video. For instance, for the song *Splitter von Granaten*, the database stores its id — *24*, the name of the song — *Splitter von Granaten*, the artist — *Adam Angst*, the genre of the artist — *Rock* and the id of the song video — *qbjaVTKEdGo*, as seen in example 95. The video thumbnail can be seen in figure 7.1.

(95) https://www.youtube.com/watch?v=**qbjaVTKEdGo**

---

[3]Adam Angst: Ja ja, ich weiß
[4]https://www.sqlite.org/index.html

Figure 7.1: An example of a video thumbnail stored in the database

Apart from that, two files in the JSON format are stored in the table as strings. One contains the time points when every line of the lyrics is sung, and the other — the lines of the lyrics. These files are used in all activities that involve the synchronisation of the video with the song text.

The table for storing constructions contains the following information: an id of a song which the constructions were extracted from, the type of the constructions — prepositional phrases, finite verb forms, verbs with prefixes or verbs in the passive voice — and a string representing a file in the JSON format. The file stores the extracted constructions, the distractors for the multiple-choice questions with these constructions, and other information needed to implement the exercises in the app. This file is used in activities that involve working with grammatical constructions in the song text.

7.3 Activities

An Android app usually consists of one or several activities. An activity includes both the appearance of the screen and the functioning of a part of the app. In this section, the app is described by discussing its separate activities and the relations between them.

7.3.1 Start activity

The search of a suitable song begins from the Start activity, shown in figure 7.2. Here, the user has a few options. He/she can either start entering an artist or a name of a song in the search field, or open a list of songs that have constructions of a certain

grammatical topic or belong to a certain genre. In the upper rectangular block, which the user can swipe to the left and to the right, the preferable grammatical topic can be selected, while the lower rectangular block contains a list of genres. By clicking on *Verb prefixes* in figure 7.2, for instance, one arrives at the list of songs that contain verbs with prefixes in their texts. After that, the user can either scroll the list to choose a song, or continue the search by typing in the search field. The options for the upper rectangular block include the four types of grammatical constructions used in the app — prepositional phrases, verbs with prefixes, finite verb forms and verb forms in the passive voice. The genres available with the current corpus of songs are hip-hop, rock, pop, singer/songwriter, electronic and metal. The list of songs the user gets by clicking on one of the genres or grammatical topics is similar to the list of search results which is shown when the user start typing a word in the search field.
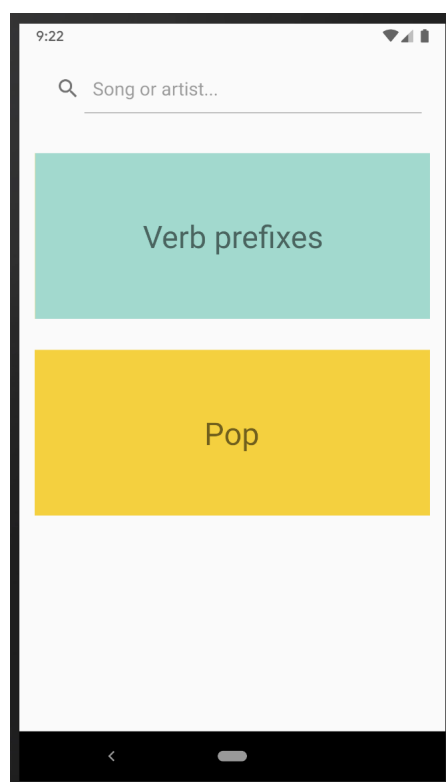
Figure 7.2: Screenshot of Start activity

The list received after typing *all* in the search field is shown in figure 7.3. The search is done using both songs' names and the names of the artists that performed
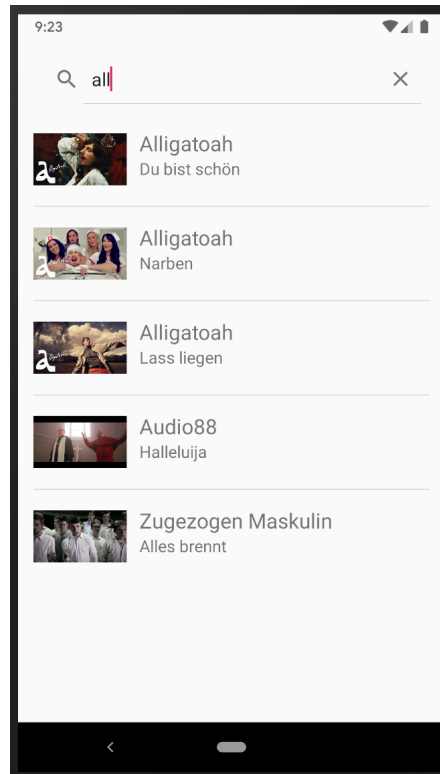
Figure 7.3: Screenshot of Start Activity. Search

them. Partial matches are allowed. By clicking on a song, the user gets redirected to the Menu activity.

7.3.2 Menu activity

As one can see from the example screenshot in figure 7.4, the Menu activity displays the thumbnail of a song's video, the artist performing the song, the song name and several buttons. These buttons stand for the modes that the user can choose for playing the song. To get redirected to the karaoke mode, one should simply press the corresponding button. While the karaoke mode does not have any additional options, the highlight and the game modes should be chosen in relation to a certain grammatical construction.

As shown in figure 7.5, when the button *Highlight Mode* has been clicked, the list of options gets expanded. The same works for the button *Game Mode*. In the options, only types of constructions found in the current song are shown. By clicking on one of the options in figure 7.5, the user gets redirected to the highlight activity where
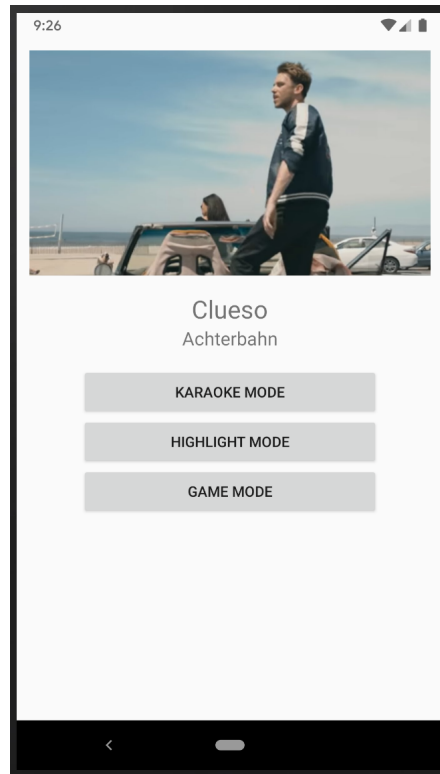
Figure 7.4: Screenshot of Menu activity. Song *Achterbahn* by Clueso

the constructions of the type specified in the option are highlighted. For instance, if the user clicks on *Verb prefixes*, the text of the song will be shown with such verbs highlighted in it.

If options for one of the buttons have already been expanded, and the user clicks on the other button with expandable options, the options for the first button get hidden.

7.3.3 Karaoke activity

When the user clicks on *Karaoke Mode* in the Menu activity, the Karaoke activity opens. In this activity, the video automatically starts playing, and the line currently sung gets highlighted, as the line with the blue background in figure 7.6. If a lyrics line is longer than the space between the left and the right side of the display, the words that cannot fit are shifted to the next text line. For instance, lyrics line *Denn ich fahr immer noch ohne Plan nach all den Jahren* is split into two text lines, as can be seen from figure 7.6. When the lyrics line gets highlighted, however, both text lines
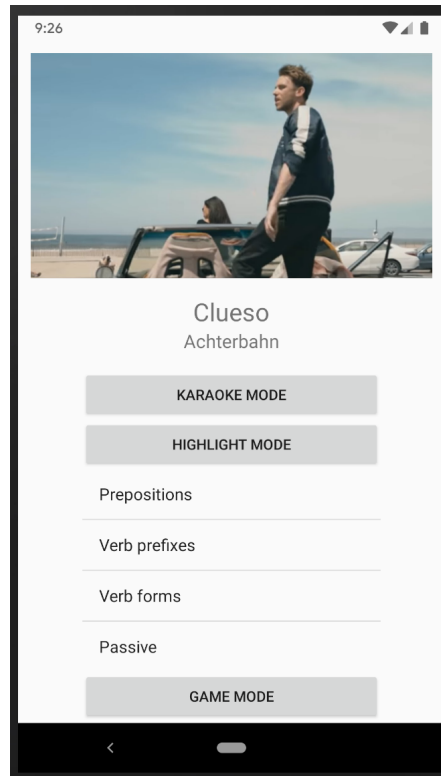
Figure 7.5: Screenshot of Menu activity with the expanded list of options for the highlight mode. Song *Achterbahn* by Clueso

are selected.

The videos are accessed with the YouTube Android Player API[5]. The opportunity to go back and further with the video controls and to pause the video is disabled, as it would hinder the synchronized playing of the song. When the video stops playing, the user is nevertheless able to choose if he/she would like to play the video again or go back to the Menu activity. See figure 7.7 for a screenshot demonstrating the options.

7.3.4 Highlight activity

The Highlight activity is rather similar to the Karaoke activity. Like in the Karaoke activity, the line currently sung gets highlighted, and the video controls are disabled. After the video has finished playing, the same options as in the Karaoke activity are shown: the user can either play the song again or go back to the menu activity.

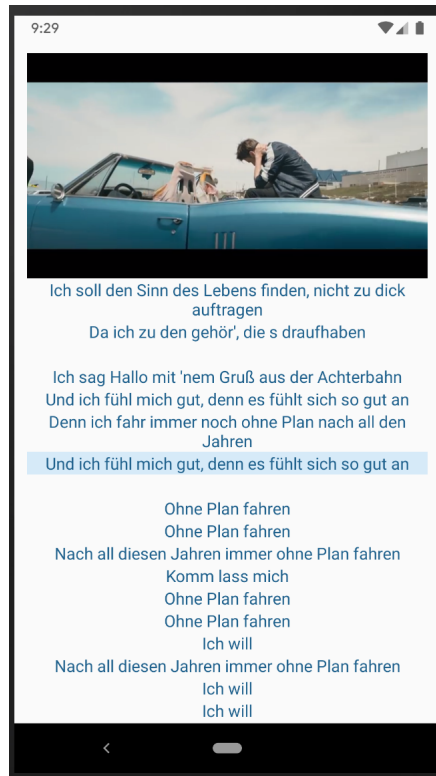[5]https://developers.google.com/youtube/android/player

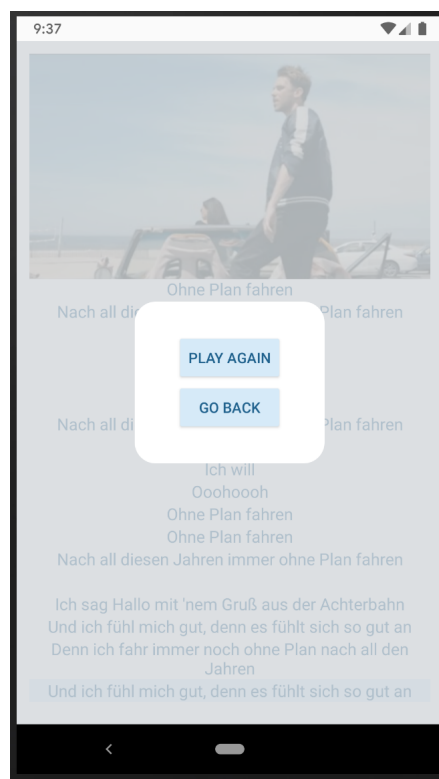Figure 7.6: Screenshot of Karaoke activity. Song *Achterbahn* by Clueso

Figure 7.7: Screenshot of Karaoke activity after the video finished playing. Song
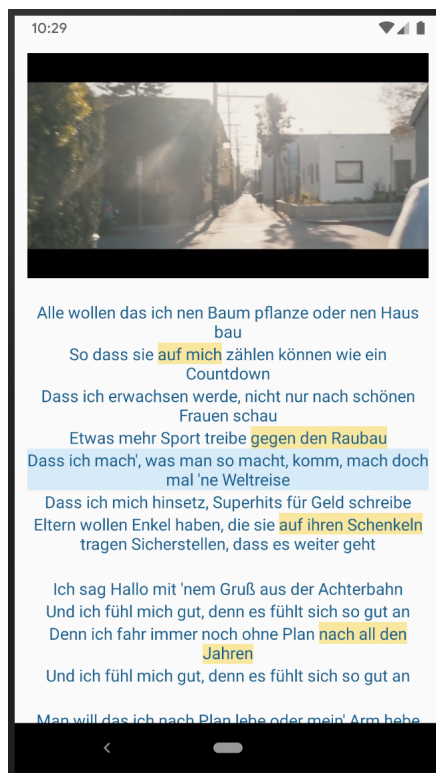*Achterbahn* by Clueso

Figure 7.8: Screenshot of Highlight activity. Song *Achterbahn* by Clueso

The only difference is the highlighted constructions shown in song texts in the Highlight activity. The type of constructions is chosen in the Menu activity. Highlighting constructions is a kind of input enhancement aimed at drawing the attention of the learner to the structure of constructions.

Different colors are used for the four types of grammatical topics. Prepositional phrases, as in figure 7.8, are marked with light-yellow, finite verb forms — with light-purple, verbs with prefixes — with aquamarine, and verbs in the passive voice — with light-red.

7.3.5 Game activity

If the user chooses the game mode in the Menu activity, he/she is redirected to the Game activity. In this activity, like in the Highlight and Karaoke activities, the song text is synchronized with the audio of the song video, and the line currently sung is highlighted. In contrast to the other activities, however, the user is supposed to answer multiple-choice questions while listening to the song.
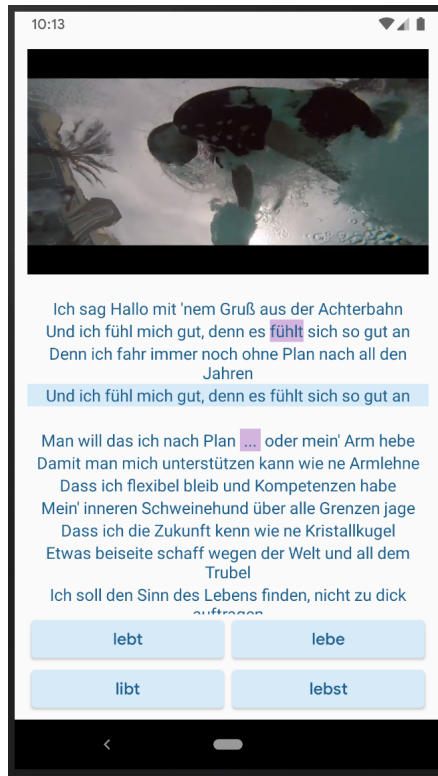
Figure 7.9: Screenshot of Game activity. Song *Achterbahn* by Clueso
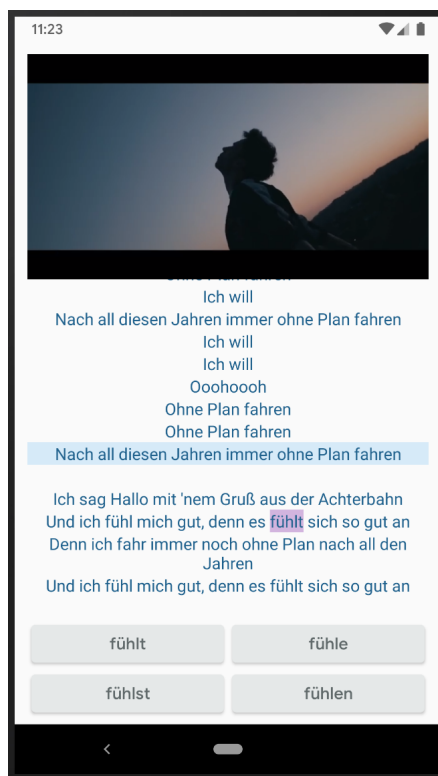

Figure 7.10: Screenshot of Game activity with disabled options. Song *Achterbahn* by Clueso

As one can see from figure 7.9, the blank is marked with an ellipsis, highlighted with the color corresponding to the construction type. Since the grammar topic chosen in the screenshot is *Verb conjugation*, the color is light-purple.

Depending on the number of distractors, there can be either two or four possible options to fill the blank with. As soon as the user clicks on the right option, the word gets inserted in the blank, as *fühlt* in figure 7.9, and options for the next blank are shown. If there are no more blanks in the song, the options for the last blank get disabled, as could be seen in figure 7.10.

For choosing the right option for the first blank in the song, the user has time from the beginning of the song. For any next blank the user is allowed to consider the options as soon as he/she answers the previous question. The user is still allowed to choose the right option as long as the line after the line with the blank is sung. If the user does not click on the right option before the second line after the line with the blank starts to be sung, the video stops playing and two options are showed, as illustrated by figure 7.11. The buttons with the answer options

Figure 7.11: Screenshot of Game activity when no right answer is given to the multiple-choice question. Song *Achterbahn* by Clueso

get disabled, and to continue playing the video, the user has to choose one of the options. If he/she chooses *Try Again*, the video starts playing from the line before the line with the blank, so that the user has one more chance to punctually click the right option. If the *Skip* button is clicked, the video continues playing from the point of time when it stopped, and the right answer is inserted in the text. The user can, therefore, solve the next blank, if there are any blanks left in the song.

When the video has finished playing, the user can either go through the quiz and play the song again or go back to the Menu activity, as in the Karaoke and the Highlight activities.

8 Summary and Outlook

To sum up, we combined the components of the successful L2 acquisition — input, attention and practice — with the grammatical aspects of German particular difficult for German L2 learners in the app for learning German. The app ensures the motivation of the learner in the process of language acquisition by including the songs popular by the majority of German L2 learners and providing the learner with an opportunity to choose his/her personalized content. Moreover, as the practical exercises should be done while listening to a song, they facilitate a high degree of the user's involvement in the process of language acquisition, which also contributes to his/her motivation.

Four different types of grammatical constructions were extracted from the song texts by utilizing the tools for the NLP analysis and applying sets of rules to the analyzed texts. The types of constructions included preposition phrases, finite verb forms, verbs with prefixes and verb forms in the passive voice. The precision of the rules varied from 98.4% to 100%, while the recall's values spanned from 76.4% to 89.2%. As the main focus was placed on the precision, the extraction of the constructions can be considered very successful.

Various types of distractors were created for the four types of constructions. For the prepositional phrases, forms of words illustrating the case that the preposition requires were used. The case, number and/or gender of forms varied. As distractors for the finite verb forms, forms of other person and/or number were used. The difference between the regular and irregular forms was also exploited. For verbs in the passive voice, wrong auxiliaries and wrong forms of suitable auxiliaries were used. Verbs with prefixes were assigned forms of verbs with the same root, but with different prefixes as distractors.

The created app has six different activities. The Karaoke activity allows the user to listen to a song and follow the song text synchronized with its audio. In the Highlight activity, a song text is enriched with highlighted grammatical constructions of the chosen type. The Game activity gives the user an access to multiple-choice questions involving the grammatical constructions found in a song

text. Other activities enable the search and the choice of the modes.

There are several strategies of improvement that can be used in the future work. First of all, a way of detecting the song texts that do not have enough grammatical constructions of a certain type can be developed. Identifying texts where some grammatical constructions are not very well distributed would also be useful. Exercises with these texts and the types of grammatical constructions can be deleted from the app, or a new type of exercises where a combination of different construction types is possible can be developed.

Apart from that, songs can be assigned language levels. For the app created as part of this project, assigning language levels was not very important, as the app was concentrated on the grammatical aspects of language, and not on its vocabulary. Being able to select a suitable language level would nevertheless make the exercises more relaxing for the learner, as he/she would not be too overwhelmed with a great number of unknown words.

Other improvements, such as correcting orthographic mistakes or splitting texts into shorter lines involve a lot of manual work. As the project was focused on the processes that could be done automatically, this kind of refinement was not considered necessary.

The last but not the least, the app can be improved by including more songs of various artists and genres. In this way, the probability of meeting the tastes of German learners of different age groups and different spheres of interest would be higher.

Bibliography

Joakim Bohnet, Bernd; Nivre. A transition-based system for joint part-of-speech tagging and labeled non-projective dependency parsing. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 1455–1465, 2012.

Sabine Buchholz. *Entwicklung einer lexikographischen Datenbank für die Verben des Deutschen*. Diplomarbeit, Universität des Saarlandes, 1996.

Robert M. Dekeyser. Introduction: Situating the concept of practice. In Robert M. Dekeyser, editor, *Practice in a Second Language*, pages 1–18. Cambridge University Press, 2007.

Zoltán Dörnyei. *Motivational Strategies in the Language Classroom*. Cambridge University Press, New York, USA, 2001.

Martin Durrell. *Hammer's German Grammar and Usage*. Routledge Reference Grammars. Routledge, sixth edition, 2017.

Georgi Dzhambazov, Ajay Srinivasamurthy, Sertan Senturk, and Xavier Serra. On the use of note onsets for improved lyrics-to-audio alignment in turkish makam music. In *Proceedings in 17th International Society for Music Information Retrieval Conference (ISMIR 2016)*, 2016.

Hiromasa Fujihara and Masataka Goto. Three techniques for improving automatic synchronization between music and lyrics: Fricative detection, filler model, and novel feature vectors for vocal activity detection. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008.

Hiromasa Fujihara and Masataka Goto. Lyrics-to-audio alignment and its application. In Meinard Mueller, Masataka Goto, and Markus Schedl, editors, *Multimodal Music Processing*, Dagstuhl Follow-Ups. Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik GmbH, Dagstuhl Publishing, Saarbruecken/Wadern, Germany, 2012. doi: 10.4230/DFU.VOL3.11041.23.

Hiromasa Fujihara, Masataka Goto, Jun Ogata, Kazunori Komatani, Ogata Tetsuya, and Hiroshi G. Okuno. Automatic synchronization between lyrics and music cd recordings based on viterbi alignment of segregated vocal signals. In *Eighth IEEE International Symposium on Multimedia (ISM'06)*, San Diego, CA, USA, 2006.

Rong Gong, Philippe Cuvillier, Nicolas Obin, and Arshia Cont. Real-time audio-to-score alignment of singing voice based on melody and lyric information. In *Interspeech 2015*, 2015.

Denny Iskandar, Ye Wang, Min-Yen Kan, and Haizhou Li. Syllabic level automatic synchronization of music signals and text lyrics. In *MM '06 Proceedings of the 14th ACM international conference on Multimedia*, 2006. URL http://dl.acm.org/citation.cfm?id=1180639.

Josef Koller, Gudrun Pfab, Jürgen Rotschedl, Monica Braun, Ulla Braune, Konrad Brandmüller, Thomas Kerscher, Barbara Moser, Sylvia Theisen, and Armin Wurpes. *MitSprache fördern: Materialien zur Sprachförderung von Schülerinnen und Schülern mit Migrationshintergrund an Gymnasien und Realschulen*. Staatsinstitut für Schulqualität und Bildungsforschung München, München, Deutschland, 2013.

Stephen Krashen. *The input hypothesis*. Longman, London, England, 1985.

Anna Marie Kruspe. *Application of Automatic Speech Recognition Technologies to Singing*. Doctoral, Technische Universitaet Ilmenau, Ilmenau, Germany, 2018.

Kyogu Lee and Markus Cremer. Segmentation-based lyrics-audio alignment using dynamic programming. In *Proceedings of the 9th International Conference of Music Information Retrieval (ISMIR 2008)*, 2008.

Alex Loscos, Pedro Cano, and Jordi Bonaca. Low-delay singing voice alignment to text. In *Proceedings of the 1999 International Computer Music Conference (ICMC 1999)*, 1999.

Matthias Mauch, Hiromasa Fujihara, and Masataka Goto. Integrating additional chord information into hmm-based lyrics-to-audio alignment. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(1):200–210, 2012. ISSN 1558-7916. doi: 10.1109/TASL.2011.2159595.

Annamaria Mesaros and Tuomas Virtanen. Automatic alignment of music audio and lyrics. In *Proceedings of the 11th International Conference on Digital Audio Effects (DAFx-08)*, 2008.

Terence Odlin. *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge Applied Linguistics. Cambridge University Press, 1989. ISBN 9780521371681. URL `https://books.google.de/books?id=7voXQ0YdJckC`.

Lourdes Ortega. *Second Language Acquisition*. Understanding Language Series. Routledge, New York, USA, 2013.

Peng Qi, Timothy Dozat, Yuhao Zhang, and Christopher D. Manning. Universal dependency parsing from scratch. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 160–170. Association for Computational Linguistics, 2018.

Akira Sasou, Masataka Goto, Satoru Hayamizu, and Kazuyo Tanaka. An autoregressive, non-stationary excited signal parameter estimation method and an evaluation of a singing-voice recognition. In *Proceedings in IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005 (ICASSP '05)*, pages 237–240. IEEE, 2005. ISBN 0-7803-8874-7. doi: 10.1109/ICASSP.2005.1415094.

Anne Schiller, Simone Teufel, and Christine Stöckert. Guidelines für das tagging deutscher textcorpora mit stts: (kleines und großes tagset), 1999.

Richard Schmidt. Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult. *Sociolinguistics and language acquisition*, 137: 174, 1983.

Richard Schmidt. Consciousness and foreign language learning: a tutorial on the role of attention and awareness in learning. In Richard Schmidt, editor, *Attention and awareness in foreign language learning*, pages 1–63. National Foreign Language Resource Center, Honolulu, USA, 1995.

Rico Sennrich, Gerold Schneider, Martin Volk, and Martin Warin. A new hybrid dependency parser for german. In Christian Chiarcos, Richard Eckart de Castilho, and Manfred Stede, editors, *Von der Form zur Bedeutung: Texte automatisch verarbeiten / From Form to Meaning: Processing Texts Automatically. Proceedings of the Biennial GSCL Conference 2009*, pages 115–124. Gunter Narr Verlag Tübingen, 2009.

Michael Sharwood Smith. Consciousness-raising and the second language learner. *Applied Linguistics*, II(2):159–168, 1981. ISSN 0142-6001. doi: 10.1093/applin/II.2. 159.

Michael Sharwood Smith. Input enhancement in instructed sla. *Studies in Second Language Acquisition*, 15(2):165–179, 1993. ISSN 0272-2631. doi: 10.1017/ S0272263100011943.

Ye Wang, Min-Yen Kan, Tin Lay Nwe, Arun Shenoy, and Jun Yin. Lyrically: Automatic synchronization of acoustic musical signals and textual lyrics. *IEEE Transactions on Audio, Speech, and Language Processing*, 16(2), 2008.

Chi Hang Wong, Wai Man Szeto, and Kin Hong Wong. Automatic lyrics alignment for cantonese popular music. *Multimedia Systems*, 12(4-5):307–323, 2007. ISSN 0942-4962. doi: 10.1007/s00530-006-0055-8.