MASTERARBEIT | MASTER'S THESIS

Titel | Title

Extramural English and lexical diversity in writing

A study of Viennese lower-secondary students

verfasst von | submitted by

Thomas Kandolf BEd

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Education (MEd)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 199 507 510 02
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Lehramt Sek (AB) Unterrichtsfach Englisch Unterrichtsfach Geographie und wirtschaftliche Bildung
Betreut von | Supervisor:	Univ.-Prof. Mag. Dr. Ute Smit

Table of contents

List of tables

List of figures

List of abbreviations

CI	confidence interval
EE	Extramural English
EFL	English as a foreign language
IDLE	informal digital learning of English
LD	lexical diversity
M	mean
MATTR	moving-average type-token ratio
Mdn	median
MTLD	measure of textual lexical diversity
OILE	online informal learning of English
TTR	type-token-ratio

1 Introduction

English surrounds us to a much greater extent than the generations before. The processes of globalisation enabled an increased spread of English across the globe and manifested its status as the global lingua franca (Lo Blanco 2014: 314). It's "inevitable presence" (Mair 2020: 13) can be felt in many areas of our everyday lives, such as politics, science, tourism, media, popular culture, and many more. Thus, the distance a learner of English experiences in relation to the target language has been dramatically reduced (Lo Blanco 2014: 323). In addition, the increased use of technology and digital devices allowed students to not solely rely on formal education to get in contact with English and its speech community, as the amount of engagement learners, especially teenagers, experience with English outside the classroom has substantially increased in recent years (e.g. De Wilde et al. 2020; Peters 2018; Lee and Sylvén 2021). Research on language learning in this field has used different terminology to describe such engagement with English away from school, such as informal language learning (e.g. Jurkovič 2019; Lee & Dressmann 2018), incidental language learning (e.g. De Wilde & Eyckmans 2017; Peters & Webb 2018), out-of-class learning (e.g. Cole & Vanderplank 2016; Peters 2018), or extramural English (e.g. Sundqvist 2009; Leona et al. 2021; Olsson 2016). In her survey of studies with a focus on English learning outside the classroom, Schwarz (2020: 22) found that the majority of studies use out-of-class or out-of-school learning to describe their focus, while extramural English was used the second most. In this thesis, the contact or involvement with English outside the classroom is referred to as extramural English (henceforth EE).

This heightened availability of EE input has led to an increase of teenagers' total time spend in contact with EE, exceeding the time of language input students receive via instructed English lessons at school (Olsson & Sylén 2015: 91; Schwarz 2020: 322). As a result, EE has developed into a critical element of learners' language development and "such informal language use [EE] may in fact provide a key support for language learning and teaching" (Schurz, Coumel & Hüttner 2022: 35). English language teachers therefore should consider their students' EE practices and acknowledge them as a vital factor of language learning.

This thesis therefore focuses on EE practices of learners of English and aims at investigating how EE influences language learning. While some studies on EE have included younger students, the majority of existing research has focused on older students – typically those in upper-secondary education or at university level. In contract, this thesis seeks to broaden the scope by examining

younger, lower-secondary students in a Viennese school who represent a diverse range of cultural, economic, and linguistic backgrounds. This demographic offers an opportunity to study how learners from varied backgrounds engage with authentic English in real-world contexts. Furthermore, most studies in this field focused on correlations between EE and vocabulary learning. However, relatively little research has been done on the benefits that frequent EE engagement might offer for other language skills, such as writing. This gap underscores the need for studies that connect EE to writing proficiency. This is why this study aims at doing just that and particularly how EE correlates to lexical diversity, as an indicator of writing proficiency.

Exploring the relationship between EE and writing among such a diverse group of participants has the potential to yield valuable insights for language teaching. It can particularly illuminate how authentic English language input students encounter outside the classroom can be used to improve learning and teaching of writing inside the classroom. Furthermore, from a general perspective on language learning, connecting those experiences students have with English outside the classroom with instructed language learning inside the classroom could also potentially increase students' motivation and overall language performance (Henry et al. 2018: 264-266).

2 Language learning outside school and extramural English

The following chapter is concerned with illustrating the concept of EE with regards to language learning, its connection to related concepts, and what correlations studies found between EE and language learning. As it is the central concept of this study, I first define EE and how it relates to language learning. Second, a model of language learning beyond the classroom as well as other important concepts will be introduced that are useful to further conceptualise EE. The last part of the chapter presents a literature review of studies on EE and language learning and connects them to the specific research context of this thesis.

2.1 Extramural English and concepts of language learning outside school

First coined by Sundqvist (2009) in her doctoral thesis, extramural English refers to English students encounter outside the walls of the classroom. Contact or engagement with English in EE is learner-initiated and voluntary, as opposed to learning initiated by other people, such as teachers, tutors, or parents. EE includes activities where learners process input, produce output, and/or engage in interactions, but these activities are not connected to any type of schoolwork, tutoring lessons, or

any other tasks connected to educational institutions, all of which are usually tied to assignments given by other people (Sundqvist and Sylvén 2016: 6-7). However, as Sundqvist and Sylvén (2016: 6) note "there is also the possibility that learners engage in specific EE activities because they feel pressured to do so, for whatever reason". An example for such a situation would be that a learner wants to watch a specific movie in English because their peers were talking about it in class. This particular learner then wants to engage in an EE activity in order to fit in and join the conversation. Nonetheless, the initiative to come in contact with English outside the classroom comes from the learner him or herself. Sundqvist and Sylvén (2016: 11) note that this engagement does not necessarily have to be active and connected to any type of action. Mere exposure to English is also regarded as an EE activity.

The question of intentionality in connection to learning is also an important one to answer when defining a model of language learning. Here, EE is an inclusive concept that involves intentional as well as unintentional or incidental learning. Thus, it encompasses all activities or situations where there is a deliberate intention to learn on the part of the learner; however, having an intention to learn is not a necessity (Sundqvist 2009: 25). Contact with English can stem, for example, from a want for entertainment without the objective to learn a language. It only matters that contact happens outside the walls of the classroom, is learner-initiated, and voluntary. As becomes clear when regarding the scope of EE in terms of intentionality, it is closely related to the concepts of incidental learning, also sometimes referred to as unintentional learning, and implicit learning (Sundqvist and Sylvén 2016: 8-9).

As already mentioned, EE is connected to many concepts that attempt to capture language learning in an informal setting outside school. Informal learning refers to learning – with or without an intention to learn – outside of organised, structured learning environments and educational institutions (Dohmen 2001: 18ff). In such informal settings, learning can be defined through various concepts, focusing on different dimensions of learning. How the most crucial ones to define EE in more detail align with and differ from EE is discussed in this chapter.

The essential defining factor of EE is that it takes place outside the classroom. To conceptualise such a setting in relation to other learning environments, Benson (2011) introduced a model of L2 English learning and teaching that he calls 'language learning beyond the classroom'(LBC). It serves as a framework for analysing how specific LBC activities are configured and how actors participate in those activities (Reinders & Benson 2017: 562). He identifies four dimensions of L2 learning and teaching environments, all of which represent an axis of a coordinate system with two extreme

endpoints: location, formality, pedagogy, and locus of control. First, **location** refers to where the learner is situated in relation to an institutionalized teacher-led setting within the walls of the classroom (situated on one extreme end of the axis). Out-of-class or out-of-school learning serves as the other extreme point and takes place outside of school but is still connected to schoolwork. In the middle are extracurricular or extramural activities that still take place at school but lack the formality of instructed classroom activities. This type of learning is usually learner initiated (Benson 2011: 9). Next, **formality** is concerned with the extent to which a learning environment is institutionalised. According to Benson (2011: 10) a formal or institutionalised setting involves professional instruction, adherence to a structured curriculum, and receiving certification upon completion. Benson (2011: 10) also distinguishes between non-formal and informal learning, the former referring to activities connected to educational institutions that do not involve assessment (e.g. extracurricular activities) and the latter not being tied to any institutions. The dimension of **pedagogy** is concerned with instruction and intentionality. It includes instructed, self-instructed, and non-instructed learning. Instructed learning is teacher-led and there is a clear intention to learn on the part of the learners. In self-instructed learning, an intent to learn is also present but the initiative to learn comes from the learner him or herself. Non-instructed, or "naturalistic" (Benson 2011: 11), learning exhibits no intentionality and no instruction. The last dimension, **locus of control** deals with "who makes the major decisions about learning and teaching – the learner or someone else?" (Benson 2011: 12). Learning can be either teacher-directed or self-directed. Schwarz (2020: 10) adds that learning can also be "non-directed, when no conscious decision to learn has been made". Locus of control clearly correlates with the dimension of pedagogy. In instructed learning, the locus of control generally lies with the teacher, whereas in self-instructed and naturalistic learning contexts, the learner him or herself is usually in control. That means that the more instructed a learning environment is, the more the locus of control shifts from the learner to the teacher.

With Benson's (2011) dimensions of LBC in mind, EE can be defined in more detail. The concept is fundamentally characterised by the location of the learner. In this thesis I define the location of EE learning as out-class-learning; however, different to Benson's suggestion (2011: 9), out-of-class learning is not connected to any type of schoolwork or activities tied to other educational institutions. Thus, EE excludes any homework assignments, private lesson, preparation for exams, or any other type of tasks assigned by teachers or other people from educational institutions. Further, the location of EE is chosen to not be specified as out-of-school. As Schwarz (2020: 16) points out, students can also engage in EE inside school as

> [they] are not exclusively places for teaching and learning, and although the core element of the EE definition is its location, it is important to recognize that even at school students have periods of leisure such as break times or free periods, in which they can engage in the same EE activities as in their homes, particularly through the use of smartphones, which give them instant access to the virtual world of the internet.

In terms of formality, EE activities are informal in nature. The setting is not institutionalised, meaning learning is not instructed, does not follow any curriculum, and does not warrant any certification. The pedagogy of EE activities is self-instructed or even non-instructed, as they are learner-led with or without an intention to learn. Lastly, the locus of control in EE lies with the learner and learning is self or non-directed. However, as outlined above, an exception would be if learners are indirectly influenced by other people to engage in a certain activity because there is pressure to do so.

LBC as a model for language learning and teaching is not without its limitations. As Benson (2011: 12) himself points out that "[t]here […] appears to be no simple relationship between the location of learning (in or out of class) and locus of control". It would be simple, if we could assume that a location inside the classroom always correlates with the teacher being the driving force behind learning. Consequently, the learner would be in control outside the classroom. However, who controls a learning situation is not as straightforward to define as it seems. For example, an activity in a group project at school can very well be self-directed by the learner, even though the assignment of the group project was most likely given by the teacher. Who the driving force behind learning in this situation might be is difficult to pinpoint. I would argue that sometimes learners and teachers share the locus of control, but then at times it lies more (or even a 100%) with the teacher or more with the learner, depending on the specific activity.

Sundqvist and Sylvén (2016: 10-14) propose a model of second language learning based on Benson's (2011) framework that aims at solving the problem of the relationship between location and locus of control and allows for a visualisation of where EE is located within the field. Language learning activities (A-H) can be placed within a coordinate system with two axes, as shown in Figure 1. Benson's (2011) dimensions of location and locus of control are the two axes which range from one extreme to the other. The x-axis represents the locus of control or who the driving force behind a learning activity is. At the left end of the axis an activity is completely initiated by another person (e.g. teacher), and on the other end it is fully initiated by the learners themselves. Consequently, EE activities then are positioned on the far-right side of the x-axis. Sundqvist and Sylvén (2016: 11) also connect the locus of control to the dimension of formality: "An activity initiated by a teacher exemplifies formal instruction and, consequently, formal learning; an activity that is learner-initiated

exemplifies informal learning". Furthermore, they link intentionality to the driving force of learning, by suggesting that intentional learning mostly takes place in other-initiated activities and incidental learning in activities initiated by the learner (Sundqvist & Sylvén 2016: 12). This suggestion might not hold true for every English learning activity, as learners can also engage in EE on their own initiative with an intention to learn. Vice versa, learning initiated by other people also potentially leads to incidental learning (Laufer & Hulstijn 2001). The location of the student when being exposed to English or engaging in an English activity is depicted by the y-axis. It reaches from the learner sitting at a desk at school to the learning being "literally as far away as possible from the desk." (Sundqvist & Sylvén 2016: 11). All activities placed above the origin would then be considered EE in terms of location. The origin of the coordinate system, where both axes meet, represents a learning situation where the learner is literally at the wall of the classroom and the locus of control lies half with the learner and half with the teacher.

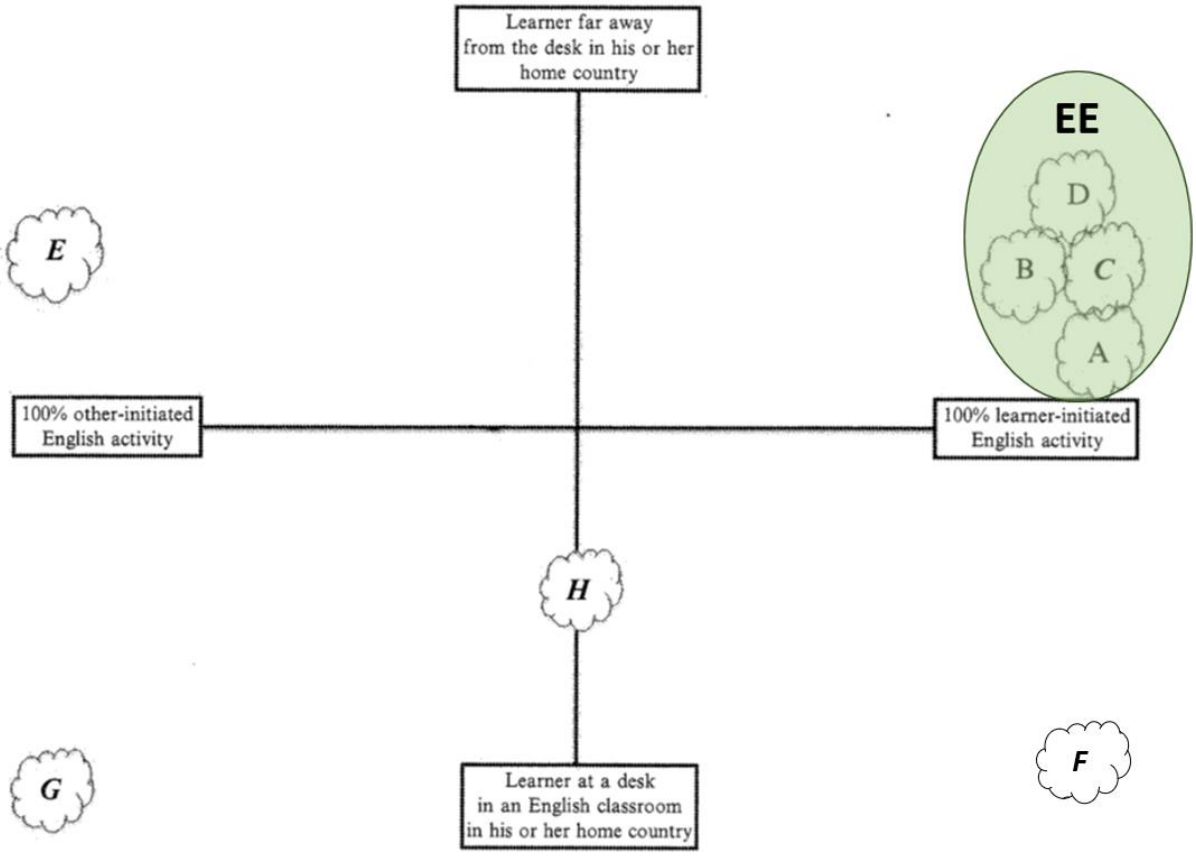

Figure 1: *Model of L2 English learning including EE (from Sundqvist & Sylven 2016: 10, own adaptation)*

Figure 1 shows extramural English (EE) as a larger area that covers different locations – from the learner being just outside the walls of the classroom to him or her being as far away from the desk in the classroom as possible – and it is located on the right extreme of the x-axis where the learner initiates contact with English. EE Activities A-D represent learner-initiated activities with varying

locations outside the classroom. The objective of the learner in these activities can either be to learn English or they can serve a different purpose such as entertainment (Sundqvist & Sylvén 2016: 12). By identifying areas rather than discrete points in the coordinate system, Sundqvist and Sylvén (2016: 13) argue that the model allows for illustrating "the dynamics of L2 English learning". Furthermore, the model resolves the issue of determining the locus of control in different locations. An activity could start at the bottom left quadrant as an activity initiated by the teacher in the classroom (G), for example watching the trailer of a series for the purpose of learning a certain language structure. This activity could then move along the x-axis where the locus of control lies with both the teacher and the learner (H), e.g. they jointly decide to watch an episode of the series at school. If the learner gets interested in the series and decides to voluntarily watch it in their spare time, the activity then moves in the upper right quadrant where EE activities are situated. The locus of control shifts from lying with the teacher, to a jointly initiated activity, to fully being with the learner. Thus, activities carried out in the classroom can develop over time into EE activities. By allowing this fluidity and by representing the driving force behind learning as a continuum, the model also solves the location-locus of control problem in Benson's (2011) LBC model. The locus of control can be shared between different parties and the model also allows for illustrating activities where the location-locus of control-relationship is contrary to what is commonly expected, e.g. learner-initiated activity in the classroom (F) or teacher-initiated activity outside the classroom (E).

There are other terms and concepts in language learning that relate to EE, such as explicit, incidental, or implicit learning. EE can be regarded as an umbrella term that includes the concepts mentioned here but goes beyond the scope of each individual one. In terms of intentionality EE includes activities with or without an intention to learn and therefore encompasses explicit (or intentional), incidental, as well as implicit learning. Sundqvist and Sylvén (2016: 8-9) suggest that these terms represent subcategories of EE or they can be regarded in linguistic terms as "co-hyponyms of EE, which is the superordinate term" (Sundqvist & Slyvén 2016: 8).

When there is an intent to learn, learning is generally understood as explicit learning (Ellis 2009 7). In EE this would entail activities outside the classroom that learners engage in to explicitly learn something that is not connected to any type of schoolwork or recommended by their teacher or other people. For example, a student intrinsically decides to improve his or her vocabulary by watching English language movies and collecting unknown words on flashcards for further revision. Incidental learning, on the other hand, refers to "learning without an intent to learn, or as learning of one thing, e.g vocabulary, when the learners' primary objective is to do something else, e.g.

communicate" (Schmidt 1994: 10). However, it is difficult to discern whether or not someone has an intention to learn. As EE includes both explicit and incidental learning, it is not necessary to make this distinction when analysing students' EE practices, but it is worth noting that these concepts are included in EE.

Implicit learning, that is learning that "takes place without either intentionality or awareness" (Ellis 2009: 7), is also clearly connected to EE. Most EE activities teenagers engage in are likely connected to entertainment, which could entail that there is no intent to learn and no or highly reduced awareness (Avello et al. 2019, Hannibal Jensen 2017, Peters 2018). There are conflicting views in the literature on whether learning is ever possible without any level of awareness. To emphasise that at least some degree of awareness is necessary in learning, Schmidt (1994: 7) argues for two types of awareness, one being noticing on a surface level and one being metalinguistic awareness of the language itself and its rules and structures. Metalinguistic awareness is essential in order for what is learnt to be assimilated into the learners' interlanguage. With this view in mind, implicit learning without any awareness does not exist, as there is always a minimal degree of awareness on the level of noticing, and "a better definition of implicit learning might be 'learning without any metalinguistic awareness'" (Ellis 2009: 7). Even though it is disputed if learning without awareness in terms of noticing is possible or not, it is widely agreed upon that implicit learning lacks awareness on a metalinguistic level (Ellis 2009: 7).

Lastly, two other concepts related to EE that are used to investigate informal learning in the digital space are informal digital learning of English (IDLE) and online informal learning of English (OILE). IDLE has been defined by Lee and Dressman (2018: 436) "as self-directed, informal digital English learning independent of formal context" and learners use "a range different digital devices (e.g. smartphones, desktop computers) and resources (e.g. web apps, social media) independent of formal contexts" (Lee 2019a: 768). IDLE does not necessitate students to be online, as it also includes using devices to learn offline. Even though the concept shares some similarities with EE, it differs in terms of location and intentionality. Besides extramural learning, it also includes extracurricular learning outside school that is still connected to school. Furthermore, Lee and Dressman's (2018: 436) definition of IDLE as "self-directed" invokes a perspective of learning as being intentional. Incidental or implicit learning do not see to be included in the concept. OILE on the other hand is conceptualised as incidental learning in a strictly online context involving "a complex range of internet-based activities" (Sockett 2014: 7). OILE is firmly linked to incidental language acquisition and therefore differs from EE which also includes intentional or explicit language learning (Sockett

2014: 14). Another difference to EE is the focus on activities taking place online and the exclusion of offline activities. IDLE and OILE show some similarities to EE and are helpful to put more emphasis on the digital space in informal learning, but these models are less inclusive in terms of intentionality and the context of learning than EE.

With this discussion of EE in mind, extramural English is defined in this study as English-language activities Viennese lower-secondary students engage in outside the classroom. Engagement with EE is voluntary, learner-initiated, and not tied to any type of schoolwork or tutoring lessons; thus, excluding assignments or activities given by the teacher or other people linked to educational institutions. Moreover, the learners can engage in EE with or without an intent to learn. Therefore, EE includes explicit, incidental, as well as implicit learning. These activities can be carried out in an online as well as offline context.

2.2 Studies on extramural English

Research on language learning outside the classroom can be traced back to the 1980s but gained real momentum only recently, as the widespread use of the internet opened more spaces for language learning to take place (Sundqvist & Sylvén 2016: 114-116). For her PhD thesis on EE, Schwarz (2020: 18-27) conducted a survey of studies investigating language learning outside the classroom. It includes 77 studies from various contexts published before 2019. Overall, there was a clear increase in research in EE since 2000. While early research efforts were mostly located in Asian contexts, research in Europe gained momentum later. The majority of studies focused on English as the target language of extramural activities, with few exceptions investigating extramural engagement with other languages, and the learners included in the studies were mostly EFL (English as a foreign language) students. In terms of the research objectives, studies included in the survey aimed at three different foci: "mapping a group of learners' out-of-class activities and/or resources […] closely describ[ing] the practices of one learner […] [and to] empirically relate extramural practices to language learning, for instance, to overall language proficiency […] or to language skills" (Schwarz 2020: 24). While a close description of learners' EE practices is not the aim of this study, it is very much located within the 1st and 3rd focus of previous research on EE. While I try to map my students' EE activities, there is also a focus on finding correlations between EE and a specific language skill, i.e. writing.

What follows is an account of studies on EE. Results of studies on EE, what EE activities learners engage in, and how EE correlated to certain language skills will be outlined. A particular focus lies on studies that link EE to writing and lexical diversity, which is the focus of this thesis.

2.2.1 Extramural English practices

When it comes to the EE practices students engage in, one could argue that they mainly engage receptively in EE for entertainment purposes. The results of the studies reviewed support this claim. Learners are mostly involved in receptive EE activities rather than productive ones (Lai, Chun & Gong 2015: 290; Schwarz 2020: 213, Sundqvist 2009: 117; Warnby 2022). Even though their ranking sometimes differs, studies reported that learners most frequently engage in EE activities, such as listening to music, watching TV/series/movies/videos, gaming, and using social media (e.g. de Wilde, Brysbeart & Eyckmans 2020; Peters 2018, Schwarz 2020, Tam & Reynolds 2022, Warnby 2022). Furthermore, the focus of common EE activities seems to be the naturalistic contact with authentic language input. In his study on the quality of IDLE activities, Lee (2019a) found that the participants engaged to a greater extent in meaning-focused activities, such as gaming or watching TV, as opposed to form-focused activities like translating an L1 text into English or practicing a specific grammatical structure. This suggests that learners prefer EE activities that entertain and where they receptively encounter authentic language input, rather than engaging in productive EE use where there is a fundamental focus on form, e.g. writing or speaking.

In line with this argumentation, several studies suggest that writing is among EE activities students engage in the least (Cabot 2016: 173; Inaba 2018: 92; Olsson & Sylvén 2015: 87). On the rare occasion that writing is mentioned in connection with popular EE activities, it only includes writing posts on social media as opposed to genres typically taught at school, such as e-mails or blog posts (Olsson & Sylvén 2015: 87). In addition to form-focused activities being unpopular, Inaba (2018: 92) suggests that the lack of conversation partners and opportunities for meaningful written interaction might be a possible reason for students not to engage in EE writing. Another EE activity rarely done by learners is reading (Avello et al. 2019: 84; de Wilde, Brysbeart & Eyckmans 2020: 177; Sundqvist 2009: 117; Warnby 2022: 131). When viewing reading as extensive reading of books or newspapers, this lack of EE engagement might be connected to the decreasing use of print media in EE. While older studies report more frequent use of traditional print media (Verspoor, Bot & van Rein 2012: 157), newer results suggest a shift towards digital media and a tendency that moves away from printed books (Peters 2018: 152; Schwarz 2020: 214; Tam & Reynolds 2022: 14). Whether e-readers have become more popular in EE due to this development is going to discussed here later. Yet, online

reading seems to play a role in EE and students especially engage in reading social media posts (Schwarz 2020: 213). Even though using social media involves reading to a certain extent, studies suggest that reading, especially extensive reading, still has a diminished role when it comes to time spent on EE.

There also seem to be differences in EE practices between different groups of learners. First, studies report differences between gender. In terms of quantity, i.e. frequency of engagement, male learners reportedly spend more time on EE than female learners (Hannibal Jensen 2017: 8; Olsson 2016: 70; Sundqvist 2009: 119). They also show higher results when it comes to frequency of EE (Olsson & Sylvén 2015: 86, Olsson 2016: 70). In terms of specific activities male and female students engage in, Hannibal Jensen (2017) and Olsson and Slyvén (2015) analysed gender specific differences in EE activities. Their results show that male students' favourite EE activity was gaming while the female participants preferred listening to music the most. Other studies report that age also plays a role in what EE activities learners engage in. Overall, older students come into contact with English outside the classroom more frequently than younger ones (Puimège & Peters 2019: 965; Warnby 2022: 131). The EE activities most frequently mentioned as popular with learners tend to be independent of age; however, there are some exceptions. Gaming seems to be favoured more by younger students. In addition, older students routinely watch English language movies or TV without subtitles, while younger students watch subtitled movies or TV shows (Bollansée & Puimège 2020: 206, de Wilde, Brysbeart & Eyckmans 2020: 177, Tam & Reynold 2022: 21). This could be due to the fact that older learners with a higher language proficiency level might be more motivated to watch movies/TV without subtitles while younger students may lack the knowledge to comprehend the input without subtitles (Tam & Reynolds 2022: 21).

As the research for this thesis is being conducted in an Austrian context and more specifically Vienna, two previous studies done with Viennese students will be highlighted here. One is a master thesis by Hahn (2017) that also has been published as an article in VIEWS (Hahn 2018), and the other a PhD thesis written by Schwarz (2020), which was already mentioned above. Both studies investigated EE of Viennese upper secondary students and how EE translates to receptive and productive vocabulary knowledge. Their results show a similar picture and are in line with research on EE in other contexts. The most popular EE activities in both studies are listening to music, watching videos online, and using social media (Hahn 2018: 10; Schwarz 2020: 213). Interestingly, Schwarz (2020: 213-214) found that those EE activities coincided with the three favoured general spare time activities. This suggests that "English is an integral part of [the students'] lives" (Schwarz

2020: 214) and how they spend their leisure time. Furthermore, both studies reveal that the learners mostly engage in receptive activities – rather than productive ones – outside the classroom that mainly involve listening to and watching audiovisual input. Even though Schwarz (2020: 213) reports that students regularly read posts on social media, the less frequent EE activities in both studies correspond to the research outlined above, namely writing, reading (e.g. books, comics, newspaper articles), and the use of print media. Another similarity to previous research are the differences between gender. Both studies concluded that male learners overall engage in EE more frequently and that boys play more video games than girls. In contrast to results of other research, however, Hahn (2018: 14) illustrates that male participants also listen to music and use social media more often than female participants. In Schwarz' study (2020: 205) reading books and using social media were EE activities female students engaged in more frequently than male students.

2.2.2 Extramural English and language skills

The second focus of this study is to find possible correlations between EE and a specific language skill, i.e. lexical diversity in writing. Multiple studies aimed at connecting EE to different aspect of language proficiency.

The majority of research involves vocabulary knowledge, and most studies here focused on receptive vocabulary knowledge assessed with various vocabulary tests (Busby 2021; Hannibal Jensen 2017; Peters 2018; Peters et al. 2019; Puimège & Peters 2019, Tam & Reynolds 2022; Warnby 2022; de Wilde, Brysbeart & Eyckmans 2020), while some studies investigated both receptive and productive vocabulary knowledge (Hahn 2017; Lee 2019a, Schwarz 2020; Sundqvist 2009; Sundqvist 2019). Overall, the main take-away from studies on receptive vocabulary is that a high quantity of EE correlates positively with receptive vocabulary knowledge (Hahn 2017: 118; Peters et al. 2019: 772; Schwarz 2020: 236; Sundqvist 2009: 155; Puimège & Peters 2019: 996). An exception would be Lee's (2019a) study on how the quality, i.e. variety of activities, and quantity of IDLE exposure influences receptive and productive vocabulary knowledge of Korean university students. He found that the quantity of activities has no direct correlation with both receptive and productive vocabulary knowledge but the quality of IDLE does (Lee 2019a: 775). However, a likely explanation of these results is that the study reports a relatively low total time spent per day on IDLE when compared to other studies. In Lee's (2019a: 772) study, nearly half of the participants spent less than 1 hour per day on IDLE. In comparison, in Schwarz' (2020: 218) study on EE, her upper secondary students engaged in EE for 3 hours and 51 minutes per day on average. Yet, it is worth noting that quantity of

EE does not necessarily translate to language learning and that the quality of EE activities might have a more significant influence on learners' language development.

This is why it is vital to investigate how specific EE activities correlate to language skills. With a focus on receptive vocabulary, several studies found that EE activities such as gaming (Busby 2021: 72; de Wilde, Brysbeart & Eyckmans 2020: 178; Hannibal Jensen 2017: 13; Puimège & Peters 2019: 966; Sundqvist 2009: 156; Sundqvist 2019: 95; Warnby 2022: 133), watching TV/movies/series/videos (Puimège & Peters 2019: 966; Sundqvist 2009: 156; Tam & Reynolds 2022: 21; Warnby 2022: 133), reading books or online content (Busby 2021: 72; Tam & Reynolds 2022: 19; Warnby 2022: 133), listening to music or other audio input (de Wilde, Brysbeart & Eyckmans 2020: 178; Warnby 2022: 133), and using social media (de Wilde, Brysbeart & Eyckmans 2020: 178) have positive effects on receptive vocabulary knowledge. All these activities offer no or relatively few chances to use English productively and students are passively exposed to EE. One important note here is that passive exposure somewhat carries connotations of receiving language input without any active characteristics. However, one could argue that processing input and understanding its meaning does involve a degree of active engagement, albeit to a lesser extent than productive ones. Schwarz (2020: 238) identified in her study several students with a high receptive vocabulary knowledge who engage in what she termed as "niche activities". These are activities that necessitate language processing on a higher, more productive level, such as rapping in English, programming, or being a member of an English-speaking sports team. The positive correlation of niche activities with higher receptive vocabulary knowledge points to the fact that even though somewhat passive exposure to EE has a positive influence, engaging in productive, intense EE activities might yield even more pronounced effects (de Wilde, Brysbeart & Eyckmans 2020: 180). Most studies link EE to receptive vocabulary knowledge and they mostly identify positive correlations to receptive EE activities. How these correlate with productive language skills, however, is something previous research investigated to a lesser extent.

This is why this thesis focuses on how EE correlates with productive language use, namely writing. Only few studies attempted to do just that (Olsson 2016; Olsson & Sylvén 2015; Sundqvist 2019; Verspoor, Bot & van Rein 2011). In the following, three studies will be discussed that have a research objective similar to my thesis and indicate potential correlations between EE and writing proficiency.

The first study is a semi-longitudinal study by Verspoor, Bot and van Rein (2011). Over one and a half years, they compared Dutch 1st and 3rd grade secondary students (13 and 15 years old) in bilingual programs, where English is used for instruction in half the subjects, and monolingual programs. One

of the schools was located in what the authors termed the "bible belt" (Verspoor, Bot & van Rein 2011: 152) of the Netherlands. Students in this school represent a small group in the Dutch educational system who, for religious reasons, lack exposure to English outside the classroom and are not allowed to listen to music or watch TV. The participants were divided into four groups: "Monolingual Non-media (Mo/Non-Media), Bilingual Education Non-media (Bi/NonMedia), Monolingual Media (Mo/Media), and Bilingual Media (Bi/Media)" (Verspoor, Bot & van Rein 2011: 151). Students filled out an EE questionnaire and produced 3 written assignments over 3 semester. The writing tasks all had personal topics close to the students' reality and got more difficult over time. Each text was scored using a holistic score from 0 to 7. The results indicate that exposure to English media outside school positively influences students' writing proficiency. The Mo/Non-Media group scored significantly lower than all the other groups. In addition, the development in writing proficiency over the course of the study correlated with EE media exposure, as the groups with media exposure improved their scores over time to a higher degree than the non-media groups. Even though the Bi/Non-Media group showed similar or at times slightly higher writing proficiency than the Mo/Media group, making up for the lack of media input through bilingual education, the Bi/Media group still outperformed all other groups (Verspoor, Bot & van Rein 2011: 162-164). It can therefore be assumed that EE exposure, even if relatively passive engagement through entertaining media, positively correlates with students' writing performance and development over time.

Another study which produced relevant results in connection to the topic of this thesis was conducted by Sundqvist (2019). She examined the effects of EE gaming on Swedish upper secondary students' vocabulary knowledge, as well as the lexical diversity (henceforth LD) and lexical sophistication in their writing. The texts collected were part of a nationwide test in English every 9th grader (15-16 years old) had to take. An EE questionnaire was used to inquire about learners' EE gaming practices and the texts were assessed via a type-token-ratio (henceforth TTR) for lexical diversity and corpus-based methods for lexical sophistication. The analysis of the participants' EE gaming and the texts show that gamers' texts exhibit a higher TTR, thus a higher lexical diversity (Sundqvist 2019: 104). However, TTR as a measure of lexical diversity is not without its flaws, which will be discussed in a later chapter. Sundqvist's results (2019: 104) also suggest that frequent gaming also leads to a higher lexical sophistication, i.e. frequency of difficult or rare words, in students' texts. Considering these results, EE gaming might positively correlate to lexical diversity and sophistication in learners' writing.

Lastly, Olsson's (2016) doctoral thesis on EE and writing proficiency offers insights into how EE correlates with lexical diversity, variation of register between two text types, as well as syntactic complexity. The thesis contains three independent studies exploring the impact of EE on students' writing. As two studies deal with academic writing of CLIL and non-CLIL students, only one study with a similar focus than this thesis will be discussed here. The participants were 16-year-old Swedish upper secondary students. To investigate their EE practices, an EE survey was conducted, and students had to keep a language diary (Olsson 2016: 62-63). The learners then had to produce two texts, a letter and a newspaper article. The text length, syntactic complexity, and lexical diversity/sophistication of these texts were then analysed. The results indicate that students with a higher quantity of EE produced texts with longer, more complex sentences and a higher lexical diversity than those students who seldomly engage in EE activities (Olsson 2016: 70). Furthermore, when comparing the two text types, students who frequently come in contact with EE used more sophisticated words in the article than in the letter. Olsson (2016: 70) concluded that this might be an indication "that register variation is greater in this group than among students with infrequent use of EE".

To conclude this literature review of EE studies, we can assert that this thesis is very much set within previous research done on EE. The foci correspond to research objectives commonly pursued by other studies. With regards to the first focus, mapping learners EE activities, studies suggest that activities for entertainment purposes, such as listening to music or watching TV/movies/series/videos are frequented the most. Less popular EE activities are reading and writing. Further, there are differences in the quantity and quality of EE in terms of gender and age. Two studies were presented that were conducted in Austria with a research context similar to this thesis. Their overall results on EE activities students engage in support previous findings. The second focus of this study is finding correlations between EE and a particular language skill. Most studies investigated receptive vocabulary knowledge and they report positive effects of EE on receptive vocabulary. While the quantity of EE influences language competence, the quality of EE activities also seems to matter. Most EE activities are receptive in character, and while these also exhibit positive correlations to language skills, productive EE engagement seems to bring about more pronounced effects on language skills. Few studies investigated EE in connection to writing. The result suggest that EE positively correlates with overall writing proficiency, LD, and syntactic complexity. Especially frequent gamers seem to be able to produce a higher LD and lexical sophistication in their texts.

This chapter provided a definition of EE and elaborated on links to related concepts and frameworks. Through this, a definition of EE for the particular research context of this thesis has been found. Moreover, various hypotheses about EE activities students engage in and how these potentially correlate with LD in writing can be drawn from the literature review of previous studies on EE. The next chapters will focus on vocabulary knowledge and lexical diversity, as this represent the other part of this thesis' research focus.

3 Vocabulary acquisition

So far the discussion revolved around the first part of the research objective in this thesis, namely EE. The second part is concerned with lexical diversity and how it is manifested in productive language use in writing. Before engaging with this topic, vocabulary knowledge acquisition, especially productive knowledge, and how it is conceptualised in the literature needs to be clarified first. In a further step, the relevance of vocabulary knowledge for writing and how it is linked to lexical diversity of texts will be highlighted.

The following sections aims at illustrating what even constitutes a 'word' and what components are included in knowing and acquiring a word. A special focus will be on productive vocabulary knowledge and what aspects of word knowledge are relevant for productive language use. These aspects are important to define in order to understand what it means for learners to have a higher or lower lexical diversity and to find a suitable counting unit for lexical items to measure lexical diversity in my data sample.

3.1 Defining lexical items

Determining what constitutes a 'word' is not as easy of an endeavour as it seems. Lexical items can be counted in different ways and can range from methods that encompass related words under one counting unit to counting every individual word as one distinct unit. How words are conceptualised in research, i.e. what counts as a word, can have great ramifications for research and pedagogy. Gardner (2007: 241) argues for the relevance of paying close attention to how lexical units are defined in research, because it greatly influences the validity of the research and its findings. In pedagogy, corpus-based vocabulary research, where word lists are based on a specific conceptualisation of lexical units, informs decisions in educational institutions, assessment, and pedagogical practices (McLean 2018: 823). Therefore, what counts as a word in research carries

assumptions about what vocabulary knowledge and abilities learners are expected to have (Brown et al. 2022: 596-597). This is why different authors argue for and against different concepts that all aim at defining what to count as a word. Here, the most commonly used concepts will be discussed, namely word families, lemmas, flemmas, types, and tokens. This section illustrates how these concepts are defined and what advantages and disadvantages they yield for research in vocabulary and LD.

The most basic way of counting lexical items are tokens and types. Obtaining a total token count of a text entails counting every individual word. For example, the previous sentence is 13 words long and consequently contains 13 tokens. Repetitions of a word, such as 'a' and 'a' in the example sentence, are counted every time as distinct tokens (Nations 2022: 8). This method is commonly used at school or university to attain a total word count of a text. While the example sentence has 13 tokens, due to the repetition of 'a', it only has 12 types. When using types for counting lexical items, we count "the number of different words" (Jarvis 2017: 538), meaning that repetitions of a word are only counted once as one distinct type. Besides tokens and types, there are other conceptualisations that aim at capturing the nature of a 'word' in a more encompassing way than simply counting individual or different words.

Word families are the prevalent unit for counting lexical items used in vocabulary research. The concept was first introduced by Bauer and Nation (1993) and encompasses the base form of a word, or head, and its inflectional and derivational forms. Bauer and Nation (1993) devised seven levels of inflectional and derivational affixes to provide guidance for teaching, learning, and research in terms of when learners can be expected to acquire which affixes. The assumption is that the more a learner progresses in knowing a specific word family the more it increases in size, as word forms from higher levels are incorporated into the learner's knowledge (McLean 2018: 824). For example, in the earlier stages of learning the verb 'to argue', the learner might be able to also connect 'argues', 'arguing', and 'argued' to the base form and might also be able to produce these forms. The affixes here represent affixes from the lower levels of Bauer and Nation's (1993) categorisation. Later, the learner progresses and acquires affixes from higher levels. More difficult forms like 'argument' or 'argumentative' then potentially increase the size of the word family. In research, forms of the verb 'argue' up to a predetermined level of Bauer and Nation's (1993) categorisation of affixes could be considered as belonging to the same word family. Words with affixes from a level higher than the predetermined threshold, such as 'argument', would then constitute a different type (Treffers-Daller & Parslow 2018: 308). There are, however, arguments against using word families as the basic

counting unit for words. Due to the assumed development of word family knowledge over time, what constitutes a word family differs from learner to learner, depending on what level of affixes he or she has acquired. This makes comparing individual learners in a study as well as comparing findings across different studies impossible (McLean 2018: 824-825). Further, Webb (2010: 507) suggests that the assumption that knowing the head form eventually leads to knowing all other members of the word family might hold true for receptive vocabulary knowledge, as recognition of other members might be easier once the head is known. Yet, for productive knowledge, learners might not be able to produce all members of the family when only the head form is known. Adopting word families for research and in education, therefore, might "overestimate […] learners' knowledge of inflectional and derivational forms" (McLean 2018: 826), which might be especially true for lower-level learners. Using word families as the basic counting unit in a study would entail the assumption that students possess the inflectional and derivational knowledge of word families, which is unlikely to hold true in the case of this study. In addition, the use of word families for productive lexical knowledge, as the case in this study, does not yield valid results for productive vocabulary knowledge of learners and is therefore not suggested by researchers (McLean 2018: 827). This is why word families are not considered as a counting unit for this study.

Two other concepts for counting lexical items are lemmas and flemmas. The lemma was first proposed by Francis and Kučera (1982). Depending on the word class, it consists of the headword, e.g. 'do', its inflections, as well as short forms (don't) (Nation 2022: 9). The English inflections are the plural, third-person singular, present and past tense, past participle, -ing form, possessive, comparative, and superlative (for the verb 'do' these would be 'does', 'did', and 'doing'). All words included in one lemma are the same part of speech, i.e. word class. This means that the verb 'walk' and the noun 'walk' are different lemmas (Nation 2022: 9). Some researchers also suggest that forms combined under one lemma must share the same meaning (Gardner 2007: 244; Stubbs 2001: 27). Using the lemma for counting lexical items in research yields some advantages. First, which forms are included in a lemma is clearer than with word families, making comparison within and across studies easier. Second, McLean (2018: 827) sees lemmas as "a pragmatic compromise when counting both receptive and productive vocabulary use, facilitating the comparison of receptive and productive studies". One criticism of the use of lemmas in linguistic computer-based research is that analysis software do not necessarily lemmatize inflected forms, instead they commonly identify them as different types. Consequently, inflected forms are registered as different lemmas, which might invalidate the use of lemma in corpus or computer-based analysis (Gardner 2007: 244-245).

However, more recently developed programmes have been improved in this regard, e.g. UDPipe, Bitext Lemmatizer, or Python Linguistic Analysis Tool. These programmes allow for a data sample to be treated in terms of lemmatisation before the analysis. Further, the same argument against the use of word families outlined above, namely that students' do not necessarily possess knowledge of inflected forms when they know the headword, can be used for lemmas and flemmas as well.

Flemmas are similar to lemmas in that they include the same word forms as lemmas but also incorporate different parts of speech (Nation 2022: 10). The verb 'walk' and the noun 'walk' mentioned above are then two different lemmas but belong to the same flemma. Because many words in English share the same form and meaning but are a different part of speech, Nation (2022: 10) argues that flemmas are a more suitable choice for research because "dealing with different members of a flemma is little different from dealing with different members of a lemma". To underline this point, he exemplifies the similar processing load when dealing with lemmas or flemmas through the examples of 'walk' (verb) and 'walking' (present participle) as forms of one lemma (verb, -ing form), and 'walking' used as a verb and as a noun in the case of the flemma. Distinguishing between these forms necessitates the same linguistic flexibility that is "at the centre of language use" (Nation 2022: 10) and therefore exhibit a similar learning burden for students. Additionally, computer-based corpus analysis, which rarely discriminate between different parts of speech, can be done more efficiently using flemmas (McLean 2018: 842).

While there are convincing arguments for adopting either of these ways of counting lexical items, I choose to use the lemma as it represent the counting unit most suitable for my research context. Even though it cannot be readily assumed that low-level learners of English possess productive knowledge of inflectional forms of a known headword, using the lemma is the most suitable choice for this thesis. First of all, there are few instances in the data sample where the participants' knowledge might be overestimated, i.e. where two words that students might have stored as two distinct words in their mental lexicon are counted as one lemma. A sizable sample of the participants' texts was analysed by hand with regards to this problem, and the analysis shows that the occurrence of such instances is rather infrequent. Most words that are relevant in this regard are regular or irregular past tense, comparatives and superlatives, but there are few occasions where, for example, the present and past tense form of a word occur in the same text. Thus, the argument that the use of lemmas potentially overestimates learners' knowledge might not be a large factor in this thesis. Second, Treffers-Daller, Parslow and Williams (2018) advocate for the use of the lemma when analysing LD of students' texts and linking LD to CEFR scores. They found evidence that when

investigating LD "inflected forms […] should indeed be considered to be tokens of [one] type [or in this case lemma]" (Treffers-Daller, Parslow & Williams 2018: 320), whereas derivations of a head should represent a distinct lemma altogether. This view is also represented in most studies on LD, as the majority of them employs the lemma as a counting unit. Thus, the choice of the lemma in this study is also relevant for reasons of comparability with other studies. Lastly, the majority of freely available software to calculate LD, and specifically the analysis tool TAALED employed in this study, use lemmas. Using types and tokens as a counting unit would entail complex pre-processing with different software.

3.2 Word knowledge

If what counts as a word is difficult to define clearly, identifying what knowing a word entails provides a similar challenge. As this study is also concerned with how productive vocabulary knowledge is manifested in the LD of students' texts, carrying out valid research in this field requires to scrutinize what aspects are involved in acquiring vocabulary and what knowledge learners should have in order to use words productively. This is why the focus of this section is to elaborate on how vocabulary knowledge can be categorised and what knowing a word entails.

Acquiring a word is a complex and multi-faceted process and there are different approaches that aim to explain what word knowledge entails. Milton & Fitzpatrick (2014: 1-11) differentiate between three approaches to characterize vocabulary knowledge, the developmental, metaphorical, and component approach. The developmental approach aims at investigating at which stages learners acquire which parts of word knowledge. Others use metaphors, such as a lexical space of vocabulary breadth, depth, and fluency, to position different levels of learners and identify differences between them in a metaphorical approach to word knowledge. As this study is not concerned with vocabulary knowledge development or comparison of learners, but rather with what students are able to produce at a specific point in time, vocabulary knowledge is characterized using the component approach which "attempts to list the different aspects of knowing a word" (Milton & Fitzpatrick 2014: 1).

Table 1: What is involved in knowing a word (Nation 2001: 27)

form	spoken	R	What does the word sound like?
		P	How is the word pronounced?
	written	R	What does the word look like?
		P	How is the word written and spelled?
	word parts	R	What parts are recognisable in this word?
		P	What word parts are needed to express the meaning?
meaning	form and meaning	R	What meaning does this form signal?
		P	What word form can be used to express this meaning?
	concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
use	grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	constraints on use (register, frequency...)	R	Where, when, and how often would we expect to meet this word?
		P	Where, when, and how often can we use this word?

R = receptive knowledge, P = productive knowledge

Nation (2001: 27) provides a comprehensive overview of the different aspects of vocabulary knowledge. Table 1 shows the three general components of knowing a word, namely form, meaning, and use. Each of them consists of different aspects that are involved in knowing a word, for example knowing the form of a word includes the spoken and written form as well as word parts. Nation also makes a distinction between the components receptive (R) and productive (P) knowledge for each of these aspects. This means that knowing, for example, the written form of a word, requires learners to not only be able to recognise the word when they read it, but they also have to be able to write and spell it in productive use. It is important to note here that, even though one might be able to use a word productively, complete knowledge of all the aspects in Table 1 cannot be readily assumed (Schmitt 2010: 82). Especially lower-level learners might struggle grasping more complex aspects like register, word parts, or collocations.

As we can see, vocabulary knowledge is multi-layered and includes a range of aspects that can be theoretically separated but in practice correlate with each other. Therefore, it is inconceivable to measure all aspect of word knowledge, as such an endeavour would be too time-consuming, require measurements not yet or insufficiently developed (e.g. measuring knowledge about a word's registers), and the interrelatedness of word knowledge aspects potentially influences cross-test results of testing different aspects of vocabulary knowledge (Schmitt 2010: 79-80). Thus, to simplify the conceptualisation of vocabulary knowledge for research and pedagogical purposes, the distinction between receptive and productive knowledge is commonly employed. More recent as well as older studies have found that leaners generally exhibit a higher receptive vocabulary knowledge than productive knowledge (e.g. Sun & Nam 2023; Laufer 2005). Yet, how much receptive vocabulary is commonly known productively is still debatable, ranging from students being able to produce only a small percentage of receptive vocabulary (Sun & Nam 2023: 2490) to a high level of productive knowledge (Melka 1997: 92). Schmitt (2010: 80) argues that these discrepancies can be explained by "the difficulties and confusion involved in dealing with the receptive/productive issue". Thus, it is vital to properly conceptualise receptive and productive vocabulary knowledge.

Generally, receptive vocabulary knowledge is linked to listening and reading while productive knowledge involves speaking and writing (Nation 2022: 52). When listening or reading, receptive language use involves meaning recognition and recall on part of the learner. This means that we use language receptively when we are exposed to input from outside and try to understand the meaning. Productive language use entails form recognition and recall, which means that we produce language in spoken or written form in order to transmit messages to other people (Nation 2022: 52). Some researchers also use the terms passive, for receptive language use, and active, for productive language use (Meara 1990). EE activities such as watching TV or reading a book can then be regarded as passive language use. This suggests that listening and reading do not involve active characteristics (Nation 2022: 52); however, it can be argued that meaning recognition and recall in this case also requires active language processing on the part of the learner. This makes the labels of passive and active language use less appropriate. Therefore, I use the definition of receptive and productive knowledge and language use outlined above that involves meaning/form recognition and recall.

Even though it is important to define what reception and production of language entails, the relationship between receptive and productive knowledge remains unclear. When can we consider a word being part of the receptive or productive vocabulary of a learner? How does a word turn from receptive to productive? To explain this relationship, Melka (1997: 101) advocates for viewing

receptive and productive knowledge as part of a continuum. In this continuum "knowledge gradually shifts from receptive mastery towards productive mastery as more is learned about the lexical item" (Schmitt 2010: 81). What is implied here is that receptive vocabulary knowledge mostly precedes productive knowledge. This might also be due to the fact that learning to use words productively is generally more difficult than receptive mastery of a word, as form recognition and recall requires more intense learning than meaning recognition and recall (Nation 2022: 56). Furthermore, some aspects of word knowledge outlined above seem to be incorporated into productive knowledge easier than other aspects. For example, Schmitt (2010: 82) suggests that his learners were generally able to produce the stem form of a word but not its derivations. This is in line with previous arguments above, namely that word parts, such as inflectional and derivational affixes, are an aspect of vocabulary knowledge that often reaches the productive level only later. The same might be true for "contextual word-knowledge aspects, like collocation and register, […] as this type of knowledge requires a great deal of exposure to acquire" (Schmitt 2010: 82).

Although the continuum perspective makes the relationship between receptive and productive knowledge easier to grasp, it is not the only concept advocated by researchers. Meara (1990: 153), for example, suggests that words can be used productively because they are activated by lexical links or associations to other words, whereas receptive vocabulary has no such links and can only be activated from the outside, i.e. we hear or see a form when listening or reading (Nation 2022: 52-53). This view can potentially explain why leaners are able to incorporate some words into their productive knowledge relatively easily, as they might have sufficient links or associations to these particular words (Schmitt 2010: 81). This study, however, does not focus on the relationship between receptive and productive vocabulary, but on how productive vocabulary knowledge manifests itself in the LD of written production. Yet, it is important to clarify what productive knowledge of words entails and that different word knowledge aspects might be acquired at different times.

So far this section provided an overview of what vocabulary knowledge entails, what we can understand under receptive and productive knowledge, and how the relationship between the two can be understood. However, the essential question of this section remains, namely what aspects of word knowledge do learners have to acquire to use words productively in writing and consequently display a higher lexical diversity? As we can deduce from the discussion above, the answer to this question is not straightforward, but Nation's (2022: 54) conceptualisation of word knowledge (see Table 1) serves as a suitable framework to show what aspects make up the productive knowledge of a word. Although it is not the focus of this thesis whether the participants have acquired all or just

some of these, one could expect that the productive knowledge of the participants probably does not cover all aspects for most words, such as constraints on use or word parts.

To illustrate what it entails to know a word, Table 2 shows all aspects of productive word knowledge exemplified by the word 'continue', which the participants encountered in lessons preceding the writing of the texts for this study.

Table 2: Productive word knowledge aspects of 'continue' (adapted from Nation 2022: 55)

For students of this study to use 'continue' productively they have to be able to…		
…spell it correctly in writing	**form**	written
…construct the word using the right word parts in their correct form, e.g. 'continued', 'continuation'		word parts
…produce the word for expressing its meaning	**meaning**	form and meaning
…produce the word in various contexts for expressing the range of meanings of 'continue', e.g. 'to go further in the same direction', 'to keep doing something'		concept and referents
…produce synonyms and opposites of 'continue', e.g 'carry on', 'stop'		associations
…use the word correctly in a sentence	**use**	grammatical functions
…produce words that frequently co-occur with it, e.g. 'for', 'with'		collocations
…use the word to suit the formality of the context		constraints on use

This section focused on vocabulary acquisition and its connection to this thesis. It gave a rationale for using lemmas as a counting unit for lexical items in the analysis of the data sample. Furthermore, what is involved in knowing a word was illustrated using Nation's (2022) conceptualisation of word knowledge. The distinction between receptive and productive vocabulary knowledge and what aspects of word knowledge are relevant for productive language use was also outlined.

4 Lexical diversity and writing

The following chapter focuses on the writing process and lexical diversity. A process approach as a model of L2 writing will be introduced with a special focus on formulation. This will be followed by, a conceptualisation of LD and its various dimensions, how these dimensions can be measured, and which measure fits the research context of this thesis. Further, this chapter aims at establishing a relation between LD and overall writing proficiency.

4.1 The process of writing

Similar to vocabulary acquisition, writing proficiency is "a multi-faceted construct, which implies that various component skills contribute to the quality of the eventual writing performance" (Schoonen 2022: 87). The most relevant component skill of writing for the focus of this thesis is the actual formulation of written text where writers make use of their productive vocabulary knowledge. The process of formulation will be discussed in more detail below. How writers approach the production of texts can be investigated in different ways that focus on the text or product itself, the writer, or the reader (Hyland 2022: 3). Even though the focus of this study lies on the end-product of the writing process when analysing the LD of texts, the collected texts represent an instrument to also investigate the effect of participants' EE practices on their LD in writing. Thus, this section concentrates on the writer and discusses writer-oriented understandings of the writing process. This focus serves to illuminate what skills and knowledge learners have to acquire in order to produce texts with a high LD and how lower-level writers differ in their approach to writing from higher-level writers.

There are different models that aim to capture how writers approach the production of texts. They put emphasis on different aspects surrounding writing, i.e. the actual process of writing, the social context, or the interaction between L1 and the target language (Hyland 2022: 12). One of the most commonly used models of writing is the process model as the one proposed by Flower and Hayes (1981). Rooted in cognitive psychology, it suggests that the writing process consists of several mental processes that are hierarchically structured and that "may occur at any time in the composing process" (Flower & Hayes 1981: 367). The process approach views writing as recursive where different components in the process can be repeated in different stages. The main processes in the model are planning, translating, and reviewing, which in turn consist of further sub-processes. For example, planning entails generating ideas and setting/organizing goals, and reviewing consists of evaluating and revising what has been written so far. Building on this process approach to writing,

Grabe and Kaplan (1996: 216) provide an overview of specific "writing skills, knowledge bases and processes" involved in writing. Their taxonomy includes categories such as text types, topics, the writer's intention and goals, linguistic knowledge, discourse knowledge, writing skills and strategies, and others. The process relevant for this thesis is the use of linguistic knowledge, more specifically productive vocabulary knowledge. At this stage of the writing process, the writer makes use of his or her linguistic knowledge to write down previously collected ideas. Flower and Hayes (1981: 373) use the term 'translate' here to highlight the process of translating meaning or ideas that "may be represented in a variety of symbol systems other than language, such as imagery or kinetic sensations". Other authors use 'formulation' synonymously with the translation process (e.g. Schoonen 2022: 94). As 'translate' evokes a perspective of writing that merely involves translation of ideas or meaning encoded in the L1 into the target language, 'formulation' is used in this thesis instead.

The formulation process is difficult to examine and may differ from writer to writer. Schoonen's (2022: 91) illustration of the formulation process provides an overview of what linguistic knowledge students need to utilize in order to formulate English texts:

> Once the writer knows what to write, the available concepts will activate the corresponding lemmas in the writer's mental lexicon and predications need to be expressed in the most appropriate grammatical structures, including processes of linearization. At this stage, the language is still 'inner speech' and needs to be converted into written language, which means that the written (correct) word forms need to be activated or selected. The writer has to make decisions about the spelling of words. Proficient writers will know how to apply phoneme-grapheme correspondence rules or will have developed ready-made word images in visual memory.

Clearly, aspects of productive vocabulary knowledge outlined above, such as the written form of a word, play a crucial role in formulation. The higher the level of productive knowledge, the easier and more effective the formulation process might be for writers. There are, however, factors that impede an efficient formulation of texts in a foreign language. One influencing is the target language. Depending on the orthography of the target language, phoneme-grapheme correspondence rules can be applied more easily or with greater difficulties. In languages with a deep orthography such as English, where the pronunciation of words is sometimes difficult to predict from the spelling and vice versa, phoneme-grapheme correspondence rules are less straightforward (Schoonen 2022: 91). Especially beginners struggle to apply them effectively and therefore rely less on phoneme-grapheme correspondence rules when formulating texts than advanced learners. This then might

not only influence the quality of their written production, e.g. correct spelling, but also affects the speed with which beginner writers are able to bring words onto the page.

The formulation stage of the writing process can also be challenging for some writers if they lack productive lexical or syntactic knowledge. Schoonen (2022: 91) uses lemma as the lexical unit writers have to activate in the formulation process. Yet, as outlined above, one can argue that especially beginners cannot sufficiently activate lemmas and their inflected form, but only distinct types or tokens. They then have to dedicate considerable resources to the formulation process and in turn pay less attention to planning or reviewing (Flower & Hayes 1981: 373; Schoonen 2022: 94). This is often the case for lower-level or younger learners and the lack of planning and reviewing can affect text-quality and set apart novice writers from advanced ones. Bereiter and Scardamalia (1987: 10) provide two models of writing that offer a potential explanation for the differences between writing proficiency of lower-level and higher-level writers, knowledge telling and knowledge transforming. Knowledge telling refers to the way beginners approach the writing process, namely telling what they know with regards to the writing assignment and spending less time on planning and reviewing. More advanced writers, in contrast, transform their knowledge as they actively reshape their ideas and written production. They constantly reflect on their goals and problems during writing and have awareness of influencing factors in writing such as audience, appropriate style, or organisation.

In addition to helping explain the difference between beginner and expert writers, Bereiter and Scardamalia (1987: 10) also illustrate with these two models how the actual formulation process takes up substantial space in knowledge telling of younger/lower-level writers. In a study on the writing process of Spanish EFL students, Roca de Larios et al. (2008) used think-aloud protocols to investigate the various processes involved in writing. Their results show that the most time was allocated to formulation. The least proficient group even devoted over 80% of time writing to formulation (Roca de Larios et al. 2008: 38). The results are in line with previous studies investigating the writing process (e.g. Sasaki 2000). In a similar but more recent study on academic writing, Wingate and Harper (2021: 8) found that, even though students in an academic context devote less time to formulating, lower-level academic writers still dedicate considerably more time to the formulation process. These similarities in results of older and more recent studies point to the fact that in writing by lower-level writers, the formulation process has a more prominent role than other processes and with that comes a focus of beginner writers on knowledge telling rather than knowledge transforming. In this regard, the productive vocabulary knowledge writers possess greatly influences not only how much time is allocated to the formulation process but also the

overall quality of the text (Kojima, In'nami & Kaneta 2022: 163; Schoone et al. 2012: 33). It can therefore be assumed that students' vocabulary use in writing is an indicator of their productive lexical knowledge and their overall writing proficiency level (Kojima & Kaneta 2022: 114). One feature of writing commonly measured to investigate writers' vocabulary use and writing proficiency is lexical complexity and in particular lexical diversity. Kojima and Kaneta (2022) examined the correlations between L2 writing performance and text features such as syntactic and lexical complexity. Their measures for lexical complexity included lexical diversity, sophistication, and density. Their results suggest that lexical complexity, including the subcategory LD, is one of the features that significantly correlates with L2 writing performance (Kojima & Kaneta 2022: 133).

4.2 Defining and measuring lexical diversity

Defining LD goes hand in hand with how we measure it. This is why the conceptualisation of LD will be discussed together with suitable measures in one section. As mentioned above, one feature of language that is considered to be a valid indicator of a person's productive vocabulary is the LD a person's speech or writing exhibits (Jarvis 2013a: 15). Furthermore, it has been shown that LD correlates with writing proficiency, as LD is one feature of text production that positively influences human judgement of the overall quality of written texts, i.e. texts with a higher rating tend to display a higher LD (Crossley et al. 2014: 10; Vögelin et al. 2019: 59). These results support the underlying assumption of this paper, namely that the vocabulary learners use in their texts, expressed through LD scores, give an indication about their overall writing proficiency which presumably correlates with their extramural English practices.

The construct of LD captures the range of words used in a text. Expressed differently, LD measures "the proportion of words in a language sample that are not repetitions of words already encountered" (Jarvis 2013b: 88). In research, studies use different terminology to describe concepts in connection to LD. Initially, lexical richness, "the overall quality of vocabulary found in a language sample" (Kojima & Kaneta 2022: 114), and LD had been used synonymously to describe the same construct (Yule 1944). However, nowadays lexical richness is seen as a superordinate term that includes LD and other lexical features (Jarvis 2013b: 95). In this sense, lexical richness encompasses not only LD, but also lexical sophistication, i.e. the use of rare or less frequent words (Jarvis 2013b: 88), and lexical density, i.e. the proportion of content words in a text (Liu & Dou 2023: 4)

The prevalent understanding of LD In quantitative linguistics is that LD represents the opposite of repetition (Jarvis 3012a: 15). However, Jarvis (2013b: 101) argues that this assumption does not

validly reflect what the construct of LD really entails. Instead of repetition, he proposes to use redundancy inversely to LD, as repetitions are a natural feature of language. "What matters more than the sheer quantity of word repetitions in a person's language use is […] to what extent those repetitions are uselessly redundant" (Jarvis 2013b: 101), meaning that they do not add informational depth to a text. Therefore, a text that exhibits a high lexical diversity consequently includes only few or no redundancies but can, according to this view on LD, include repetitions. Jarvis (2013b: 92) sees the perspective of LD with its counterpart repetitions as the source of the problem quantitative linguistics has in this field of research. As repetition reflects the ratio between types and tokens, linguists set out to find and develop a measure that effectively captures this ratio. The challenge was and still is to find an index that is independent of text length, as longer texts tend to appear less lexically diverse using the TTR. Jarvis (2013b: 95) argues that while measures of LD can predict language proficiency, they primarily reflect the TTR rather than true lexical diversity. This focus on effective LD measures has led researchers to overlook the broader construct of LD, which lacks validity as it only captures word variety without encompassing all aspects of LD

Thus, Jarvis (2013b) proposes a multidimensional definition of LD and in a different study, aims at finding suitable ways of measuring its different aspects (Jarvis 2013a). Drawing from ecologists' definition of biodiversity and human judgements of lexical diversity, he understands LD as being more complex in its composition and derives six different properties of LD (Jarvis 2013b: 101-102): variability, volume, evenness, rarity, dispersion, and disparity. These properties of LD will be sketched in the following section.

The definition of LD given above, i.e. that LD represents the variety of words used in a text, captures the first property, namely **variability**. As already mentioned, this is the predominant focus of previous research on LD. Jarvis (2013a: 28) recommends using the measure of textual lexical diversity (henceforth MTLD) first introduced by McCarthy (2005), as it is shown to capture variability independent of text length (Zenker & Kyle 2021: 13). The second property of LD is the **volume** of a language sample. Jarvis (2013a: 28) found that volume significantly "contributes to the perceived lexical diversity of a language sample" in human judgement of LD. The measure of volume proposed by Jarvis (2013a:28 is at first sight rather straightforward, namely the total word count. However, as already outlined above, the challenge is to determine what even counts as a word. Thirdly, the distribution of tokens in relation to types is referred to as **evenness**. Jarvis (2013a: 29) advocates for using the standard deviation to measure evenness; however, he acknowledges that volume might negatively affect the results of this measure. Fourth, he also includes **rarity** in his definition of LD.

This property is commonly termed lexical sophistication and is normally not incorporated into the construct of LD, but he argues that features of language under the umbrella term of lexical richness are closely linked and have to be investigated collectively. The most common measure for rarity, or lexical sophistication, is to determine the frequency of words in a text using a corpus as a source of reference. Fifth, one important property of diversity suggested by ecologist is how well different species are distributed in space from each other. This is called **dispersion**. In linguistics, dispersion entails "the degree to which tokens of each type are dispersed evenly throughout the domain (e.g., text) as opposed to being clustered in close proximity to other tokens of the same type" (Jarvis 2013a: 25). Higher lexical diversity requires types to not appear close together but to be scattered evenly across the text. To measure dispersion, one can calculate the aggregated value of the mean distance between different tokens of the same type (Jarvis 2013a: 30). The last property of LD Jarvis (2013a: 25) proposes is **disparity**. He illustrates this by quoting Gould (1989: 49) describing disparity in relation to biodiversity: "Three blind mice of differing species do not make a diverse fauna, but an elephant, a tree, and an ant do – even though each assemblage contains just three species". In lexical diversity, this would mean that for a text to have a high level of dispersion, lexical types need to semantically differentiate to a high degree. Measuring disparity is difficult to do, as investigating semantic relationships and differentiating between them is a complex endeavour. Jarvis (2013a: 30) suggests designing a tool using the WordNet semantic sense index that calculates "the mean number of words in a text that share the same semantic sense".

To arrive at a valid score of lexical diversity, Jarvis (2013a: 31) emphasises that all of these measures of the different properties of LD have to be combined, and human judgements of how these properties affect the perceived LD of text have to be used to inquire which properties significantly correlate with LD. For the latter, Jarvis (2013a: 27) already conducted a study that showed that variability, volume, and rarity of words in a text have the strongest influence on how lexically diverse the text is perceived to be. This would mean that the results of measuring these properties should be taken into account to a greater extent when calculating LD.

As already mentioned, the greatest challenge in finding a suitable measure of LD is that it should be independent of text length. The connection between LD and sample size is best captured by Heaps' law (1978) which states that it is gradually less likely for new types to appear the longer the text is. That means that new types can be found at the beginning of a text, and with increasing length it is more likely to find tokens of types already used before. This is why a simple type-token ratio, where the total number of types is divided by the total number of tokens, is influenced by text length and

longer texts appear to be less divers than shorter ones (McCarthy & Jarvis 2007: 460). However, there are other measures that use TTR that seem to avoid sample size effects and provide a more valid LD-score, as they split the text in different segments and calculate LD scores from the mean scores of these segments (Jarvis 2013b: 94, Zenker & Kyle 2021: 13). Two of these measures will be discussed here: the measure of textual lexical diversity (MTLD) and the moving-average type-token ratio (henceforth MATTR). The MTLD uses a predetermined TTR value that serves as a threshold for the calculation. Then it calculates the TTR for the first string of words of the text until the threshold is reached. It then goes on to the next words and moves through the text until the threshold is reached again. The process is repeated until the text has been analysed back and forth. The final score is the average length of all the segments (McCarthy 2005; Zenker & Kyle 2021: 5). Covington and McFall (2010) on the other hand advocate for using the MATTR to calculate LD. The first step is to determine a window length (X) for which the TTR is calculated. The range of this window depends on the text length. MATTR computes the TTR for words 1 to X, then 2 to X+1, 3 to X+3, until the end of the text, and "the mean of all these TTRs is a measure of the lexical diversity of the entire text" (Covington & McFall 2010: 96). Both measures provide a score between 0 (only one lexical item used in the text) to 1 (all lexical items in the text are different). To find out which LD index is independent of text length, Zenker and Kyle (2021) conducted a study in which LD indices were used for different text lengths. Nine LD indices were chosen, including TTR, MTLD, and MATTR. They found that both MTLD and MATTR show the highest level of independence from sample size. Further, they established that MATTR is the most stable index for texts ranging from 50 to 200 words.

This chapter examined the writing process and its connection to LD, focusing on how LD relates to writing proficiency. A process-oriented model of writing is introduced, emphasizing formulation—the stage where writers utilize their productive vocabulary knowledge to create written text. Formulation, influenced by productive vocabulary and linguistic knowledge, is identified as a significant factor in determining the quality and efficiency of written text. Furthermore, the definition and suitable measurements of LD were illustrated. LD was conceptualised as the variety of words used in a text and is seen as the opposite of redundancy. Traditional measures of LD, like the type-token ratio (TTR), often fail to account for text length and oversimplify LD as the inverse of repetition. To address these shortcomings, Jarvis (2013a, 2013b) proposes a multidimensional framework for LD, incorporating six properties: variability, volume, evenness, rarity, dispersion, and disparity. These properties collectively capture different aspects of LD, such as word variety,

distribution, and semantic richness. Finally, an argument for using the MATTR index as a measure of LD in this study was brought forth.

5 Study design and methodology

This chapter outlines how the present study is designed. It details the main aims, research questions and hypotheses this thesis revolves around. After describing the study's participants, the chapter focuses on the instruments for data collection, the data collection and preparation procedures, and what aspects the data analysis entails.

5.1 Research aims

Overall, this thesis aims to investigate possible effects of EE practices on students' writing, specifically LD. Thus, it aims to increase understanding of what EE practices learners of English as a foreign language engage in and how these correlate with productive vocabulary use in writing. To do this, two aims will be pursued.

The first aim of this thesis is to establish what EE practices the participants engage in. This is not only necessary to find potential correlations to LD, but it also allows us to familiarize ourselves with what the learners do in their free time and, if at all, what relevance EE has in this context (Sockett 2014: 32). The mapping of EE activities in this thesis includes identifying quantity (frequency of EE) and quality (variety of EE activities) of EE, working out potential differences between gender, grouping participants according to EE engagement, and finding niche activities that only few participants engage in. Niche activities in this present study include activities that are done regularly by a very small number of participants and often require a higher level of language processing (Schwarz 2020: 238).

The second aim is to explore how lexically diverse the participants' written production is through calculating the MATTR of students' texts. Incorporating a measure of a specific language skill is a common procedure in studies on EE and LD can be used as a potential indicator of learners' writing skills. While the majority of research on EE focuses on receptive vocabulary knowledge, few studies explore productive vocabulary application in writing in connection to EE. Therefore, this study tries to contribute to the rather small body of research on EE and writing. Further, there are, to my knowledge, no studies on EE and writing with students as young as the participants in this study.

Following these two foci allows for discovering potential correlations between different patterns of EE engagement and LD scores of students' texts. Finding out whether English encountered outside the classroom influences learners' ability to produce lexically diverse texts represents the major aim and is tied to the general aim of every study in L2 learning and teaching, to further improve practices in this field. Naturally, this study attempts to do the same by gaining insights into what learners do with English outside the classroom and deducing possible implication for teaching practices inside the classroom.

From the aims stated above, several research questions and hypothesis can be deduced. The overarching research question (RQ) is:

> ***How do extramural English practices of Viennese (lower-secondary) 6th grade students correlate with the lexical diversity of their writing performance?***

The underlying hypothesis on which this thesis is based on is that a high quantity and quality of EE correlates positively with LD in writing. Yet, to explore whether this hypothesis proves to be true or not, several more specific research questions have to be answered first.

The first set of sub-questions is concerned with the EE practices of the participants:

RQ 1: *What extramural English activities do Viennese (lower-secondary) 6th grade students engage in?*

RQ 2: *What frequency of EE engagement do the participants report?*

RQ 3: *What are the gender differences in terms of quantity of EE?*

RQ 4: *What are the gender differences in terms of quality of EE?*

RQ 5: *What 'niche' activities do the participants engage in?*

Several hypotheses with regard to these questions can be derived from the literature review of studies on EE. Overall, it will be interesting to investigate whether the EE activities correspond with students' general leisure time activities. In terms of quality of EE, studies suggest that learners engage more frequently in receptive EE activities rather than productive activities that require more intense language processing. When it comes to specific activities, younger learners, such as the participants in this study, tend to be more involved in EE gaming than older learners. Furthermore, older learners commonly watch English language movies without subtitles while younger learners frequently use subtitles. Concerning the quantity of EE, younger learners seem to be less frequently

engaged in EE than older ones. There are also differences in the quality and quantity of EE between genders. Male learners commonly report a higher quantity of EE engagement than female learners. They also differ in terms of the specific activities they favour, with EE gaming being more popular with male learners and listening to music preferred by female learners. It is difficult to hypothesise about specific 'niche' activities the participants engage in, as they represent special, uncommon activities that may differ depending on the context the study is set in. Thus, it is not feasible to deduce potential 'niche' activities from previous studies.

The second set of sub-questions focuses on LD and its possible correlation with EE:

RQ 6: What is the LD, expressed through a MATTR score, of Viennese (lower-secondary) 6th grade students' text?

RQ 7: *How does the quantity of EE engagement correlate with LD scores?*

RQ 8: *How do specific EE activities correlate with LD scores?*

RQ 9: *How do niche activities correlate with LD scores?*

Judging from previous studies it can be assumed that productive activities that require a high level of productive language processing correlate with high lexical diversity in writing. Yet, frequent receptive engagement with EE might also lead to high LD scores. The overall hypothesis for this set of questions is therefore, that a high quantity of EE leads to high LD scores but the quality of EE activities, especially 'niche' activities, trumps quantity.

5.2 Participants

The participants in this study are all students at the same junior high school (Mittelschule) in the 10th district of Vienna. They all attend the 6th grade or second year of lower-secondary education. The participants' age ranges from 11 to 14 years. After attending four years in the 'Mittelschule', students either continue their education in a higher general secondary school (Gymansium) or vocational school (HAS/HAK), or they finish their compulsory education of nine years at the junior high school or in a polytechnical school and enter an apprenticeship afterwards.

It is difficult to determine how much prior English instruction the students in this study received before entering lower-secondary education and what language level they have attained. The curriculum of primary school in Austria dictates that in 1st and 2nd grade a total of 64 hours per year of English should be taught integrated in other subjects. In 3rd and 4th grade, learners receive one

hour of English instruction per week, which amounts to a total of around 70 lessons in those two years (BMBWF 2024a). However, some students in this study did not attend primary school in Austria. In terms of language level, the participants vary greatly. Students are divided into two groups, higher level students in the 'Standard AHS' group, which means they should be at the level of students in higher general secondary school (AHS or Gymnasium), and lower-level students in the 'Standard' group. Therefore, the range of language competence is substantially higher in 'Mittelschulen' in Austria than in 'AHS'. Yet, as all of them finished the 5th grade of lower-secondary school, according to the curriculum (BMBWF 2024b), they should all at least have reached a language level of A1 as defined in the *Common European Framework of Reference* (CEFR 2020).

5.3 Instruments

To obtain data and achieve the objectives set at the beginning of this study, two instruments were used. The EE questionnaires serves to inquire about participants' engagement with English outside the classroom and the two texts each participant wrote were used as an indicator of written LD. The EE questionnaire and the writing prompts for both texts can be found in the appendix.

The EE questionnaire was adapted from the one Schwarz used in her PhD thesis (2020). While it was already tried and tested, the questionnaire had to be adapted to suit the context of this thesis. As the participants in this study are younger than the upper secondary students in Schwarz' study, the initial questionnaire had to be shortened to avoid overload. While some items were irrelevant for this study, others which inquire about similar activities were merged. Furthermore, the questionnaire was re-formulated in an easier to understand language. Before conducting data collection in class, the questionnaire was tried out with students of the same age group but from different schools. A total of 5 students participated in the pilot, with ages ranging from 10 to 13. They read and answered the questions swiftly and had no substantial comprehension problems. It took them 15 to 20 minutes to complete the questionnaire. Yet, the students indicated some areas for improvement. Changes had to be made to the wording of the instructions and the layout in order to clarify what is asked of participants in each section. Additionally, the introductory text at the beginning of the questionnaire was reformulated to emphasize more strongly the importance of answering the questions honestly.

The adapted questionnaire consists of 4 sections that focus on different aspects. Section 1 has a total of 7 items which inquire about demographic and language background of the participants as well as which group they belong to in English at school ('AHS' or 'Standard'). In section 2, participants

indicate their access to media and the internet at home. Section 3 focuses on general leisure time activities. The last and most extensive section comprises questions about the participants' EE activities. A total of 49 items inquire about different English language activities and media types. Most of the items are multiple choice items covering general leisure time and EE activities asking students to report their engagement on a 5-point Likert scale ranging from (almost) never to (almost) daily. In addition, the questionnaire contains some open questions when additional information is required, e.g. language of subtitles when watching movies. Each student labeled the questionnaire with his or her code consisting of the second and third letter of their first name and their month of birth (e.g. 05 for May). The same was done for the two texts to ensure that a participant's questionnaire could be linked to the corresponding texts.

The two texts cover two different genres, with one being a narrative text and the other a descriptive text. The narrative text is about a family driving to their holiday destination through bad weather. The descriptive text is concerned with students' favourite sport. The rationale for collecting two text of different genres per participant is that such an approach yields a higher potential for students to show a wide variety of lexis. As the chosen measure of LD, i.e. MATTR, is independent of text length, the produced text did not necessarily have to be the same length. Nevertheless, it was decided to give a rather narrow word count of 100 to 120 words in order to potentially obtain texts of a similar length.

5.4 Data collection

Choosing this sample of participants was mostly based on considerations of easy access. As I was the English teacher of two of the four classes and I taught different subjects in the other two, conducting the study was easier than including students from a different school as well. Before conducting the study, consent from the head teacher had to be obtained, which was given in April 2023. After that, the participants' parents were contacted to seek their consent. By the end of April 2023, all parents consented to the study and to using the data for research purposes. The participants were guaranteed that the data would be anonymised and kept confidential. They also were assured that they could choose to not participate in the study at any point. Overall, there were 97 students in all four classes. Six of them are students with special educational needs or 'Inklusionsschüler*innen'. Even though they participated in filling out the questionnaire and writing the two texts, they were not included in the data sample. That left 91 eligible participants. However, 17 of them had to be excluded as they, due to absences, either did not fill out the questionnaire or

did not write one of the two texts. This reduced the actual sample size to 74 students who filled out a questionnaire and wrote two texts. 36 of them identify as female and 38 as male.

The data collection was conducted in Mai and June 2023. To avoid time in class being dedicated to this study, data collection of the questionnaire took place outside regular school lessons and all students consented to staying at school for an additional 45 minutes to fill out the questionnaire. As the texts were part of the curriculum at this time in the school year, participants wrote them during English lessons. The questionnaire was filled out by the participants and collected from May 2nd to May 5th. Even though the students knew what the study was about, I reiterated the purpose of it and emphasized that they were not going to receive marks and should answer the questions honestly. The different sections and question formats were explained to the students before they started filling it out. As for the texts, writing prompts were designed together with a colleague and students received instruction on the topic a couple of lessons beforehand as part of the regular curriculum for this grade. The narrative text was written in all four classes between May 15th and May 18th. Before writing students revisited vocabulary related to weather, watched a video of a live weather forecast in bad weather, and talked about experiences with bad weather. The participants wrote the second text about their favourite sport between May 22nd and May 25th. In order to activate students' background knowledge, I again did some pre-writing activities with them where they brainstormed about their favourite sport, gave reason why it was their favourite sport, and collected useful words to describe it. The texts differed considerable in terms of length, as some students had difficulties sticking to the given word count.

5.5 Data preparation

Before conducting the analysis, the gathered data from the instruments described above had to be prepared for analysis. For the EE questionnaire this included coding the different items, creating summary variables, and entering the data into excel. The texts had to be transcribed and converted into txt. files in order to analyse them with TAALED (Kyle, Crosley & Jarvis 2021), an analysis tool that calculates, among others, the MATTR lexical diversity index of texts.

Each questionnaire item was assigned a code and the data from the questionnaires was entered by hand into an excel file. Data entry for the multiple-choice items and 5-point Likert scale items was straightforward. For open questions, same or similar answers were given the same code. Furthermore, a total of six summary variable were created[1]. Each variable is made up of items in the

[1] A list of all summary variables and the items they include can be found in the appendix.

questionnaire that target a similar EE activity and the highest value indicated was used as the value for the summary variable. For example, item 4c "watching English language movies", which includes watching movies on TV, on the internet with or without subtitles, and in the cinema, represents the summary variable "watching movies". To obtain an overall value about the frequency of EE engagement for each participant, an EE median scores was calculated from of all values concerning EE activities. It can range from 5 (daily engagement) to 1 (no engagement). As most of the data from the questionnaire are ordinal in nature, the median was used instead of the mean. Even though an EE median score is a convenient way to compare the frequency of EE engagement of participants, scores from individual EE activities should still be considered in the analysis as a median score might not be a reliable representation of EE engagement.

The preparation of the texts for analysis involved transcription and calculation of LD scores. Before transcribing the texts, detailed criteria for the transcription process had to be established in order to obtain an accurate measure of LD. The criteria used in this study are based on previous studies on LD (Bulte & Housen 2014; Maamuujav, Olson & Chung 2021; Wang 2014; Yu 2010) and on how TAALED conducts text analysis. One of the main issues was how to deal with spelling mistakes and whether to correct or exclude words in such instances. Most studies corrected spelling mistakes (Bulte & Housen 2014; Crossley, Salsbury & McNamara 2012; González 2017; Maamuujav, Olson & Chung 2021; Vögelin et al. 2019; de Wilde 2023). They argue that spelling mistakes do not necessarily indicate missing command over productive vocabulary, i.e. constituting an error, but could very well be a typo, i.e. a mistake. However, it can be argued that spelling is part of productive vocabulary knowledge and is therefore one indicator of the diversity of student's lexical repertoire. Therefore, a balance between correction and exclusion of words had to be struck. In this study, misspelled words were corrected if they were written correctly elsewhere in the text, if one letter was missing, if one letter was falsely added, and if one letter was misplaced within the word. Misspelled words that did not meet these criteria were excluded. This should to a certain extent account for typos without disregarding a potential lack of productive vocabulary knowledge. Furthermore, names were also excluded as TAALED does not register them as distinct words. The same applies to abbreviations such as 'can't' or 'don't', which were spelled out in the transcription. The transcription was done by hand.

As outlined above, measuring LD is not straightforward, as different aspects influence how lexically diverse a text is perceived to be. Measuring each individual property, however, might not be feasible for this thesis. Given that variability, the property most studies used to conceptualise LD, is one of

the aspects with the greatest effect on perceived LD, and that studies have shown a direct correlation between variability and writing proficiency (Crossley et al. 2014: 10; Vögelin et al. 2019: 59), a measure that aims at capturing this property of LD will be used in this study. This is why the MATTR index was employed to attain LD scores for the students' texts. However, it is important to note that this does not fully capture the construct of lexical diversity outlined above and that arriving at all-encompassing score of LD, more measures might have to be used.

Three different LD scores were calculated for each participant using the MATTR index. One score for each individual text and a mean score of both texts. As with the EE median score, the LD mean score is useful as an overall indicator of LD and to compare participants. However, as the two texts are of different genres, using the separate score of each text in the analysis might also yield interesting results. This is why the analysis does not solely rely on the LD mean score.

5.6 Data analysis

The statistical data analysis involved running tests to investigate differences between certain groups and potential correlations between LD and different EE variables. It was carried out using JASP (JASP Team 2024), an open-source and freely available statistics programme.

To statistically examine differences between certain groups, instead of the commonly employed t-test, this study used the Mann-Whitney u-test [2]. Both tests are based on specific assumptions about the data under scrutiny. The t-test as a parametric test assumes, for example, normally distributed data while the non-parametric Mann-Whitney u-test does not require normally distributed data. To examine whether the data of this study are normally distributed, Scatterplots as well as the Shapiro-Wilk test were used. The scatterplots show slight curves, which should not be the case in normally distributed data. The Shapiro-Wilk test further confirms that the data in this study are non-normally distributed[3]. Therefore, the non-parametric Mann-Whitney u-test was used.

Results of tests such as the Mann-Whitney u-test or the Shapiro-Wilk test are commonly interpreted using the p-value. Researchers can set an alpha-level where results with a smaller p-value are regarded as statistically significant. The most frequent alpha-levels in research are 0.05, used in this study, 0.01, and 0.001 (Cohen, Manion & Morrison 2018: 740). However, using the p-value as a sole indicator for interpreting results is not without its limitations. The p-value is influenced by sample

[2] Also sometimes referred to as the Wilcoxon rank sum test.
[3] A p-value smaller than .05 means that the distribution of the sample data is significantly different from a normally distributed sample.

size, which means that "any mean difference between groups or any correlation will reach statistical significance, given a large enough sample" (Plonsky & Oswald 2014: 879). Furthermore, it invites dichotomous classification of results as significant or not significant and the set alpha-levels are often quite arbitrary boundaries of statistical significance (Plonsky & Oswald 2014: 880).

It is therefore proposed to not only rely on p-value when reporting results, but to also include the effect size, which represents "an objective and […] standardized measure of the magnitude of [an] observed effect" (Field 2012: 57). One such standardized measures is Cohen's d. To interpret the size of an effect, Cohen (2013: 25-26) proposed to view a d-value of .200 as a small effect, .500 as a medium effect, and .800 as a large effect. However, he also emphasized that such benchmarks should be field specific (Cohen 2013: 25). Thus, the following L2 research specific benchmarks advocated by Plonsky and Oswald (2014: 889) are being used in this study:

$.000 \leq d \leq .400$	small effect size
$.400 \leq d \leq .700$	medium effect size
$.700 \leq d \leq 1.000$	large effect size

To make the interpretation of the statistical analysis more reliable, this study also reports the 95% confidence interval (henceforth CI) in addition the p-value and effect size. The CI exhibits a lower and an upper boundary in which centre the mean is located. For a 95% CI this means that 95% of the time a study is replicated, "the true value of the population mean will fall within these limits [or boundaries]" (Field 2012: 45). A small or narrow interval indicates that the mean in a statistical analysis is close to the true mean, while a wide interval points to the sample being "very different from the true mean, indicating that it is a bad representation of the population" (Field 2012: 46).

Due to the data not being normally distributed and ordinal in nature, a non-parametric correlation coefficient has to be employed in this study. Correlation coefficients are essentially effect sizes that tell us the magnitude and direction (positive or negative) of a correlation. Spearman's rho (ρ) and Kendall's tau (τ) are examples of such non-parametric correlation coefficients. As the data sample in this study is rather small and exhibits many tied ranks, i.e. in ranked scores many ranks have the same score or value, it is suggested to use Kendall's tau instead of Spearman's rho (Field 2012: 225). Even though Kendall's tau yields more conservative results (lower effect sizes), it allows, in the case of the present data sample, for more accurate interpretations than Spearman's rho (Field 2012: 225). Plonsky and Oswald (2014: 889) also provide suggested benchmarks for correlation coefficients L2

research. In the case of Kendall's tau, τ close to .250 should be considered a small effect, .400 medium, and .600 large.

6 Analysis and results

The following chapter states the results of the quantitative data analysis. First, data collected via the questionnaire is going to be presented. This includes background information about the participants, their general leisure time activities as well as EE practices. Next, the LD scores of students' texts are going to be discussed before delving into how these scores correlate with EE.

6.1 Participant's background

Besides EE, the questionnaire also required participants to give additional information about themselves. This allows for further characterizing the participants in this study and with that sharpening the context in which this study is set. This section presents students' demographic information, language background, and to what media they have access. Further, it contains information about their general leisure time activities.

6.1.1 Demographic information and linguistic background

As already mentioned above, there are 74 students participating in this study with a varying demographic and linguistic background. Table 3 depicts the gender, year of birth, and place of birth of all participants. The 36 female and 38 male learners are born between 2011 and 2009. They come from 11 different countries. The majority (77.03%) is born in Austria, but many of those students have a migration background, with the generation of their parents or grandparents having migrated from other countries to Austria. This cultural diversity is especially evident when looking at the languages students use at home.

Table 3: Demographic information

gender	N	%
female	36	48.65
male	38	51.35
year of birth	N	%
2011	36	48.65
2010	32	43.24
2009	6	8.11
place of birth	N	%
Austria	57	77.03
Syria	4	5.41
Iran	3	4.05
Serbia	2	2.70
Chechia	2	2.70
Sudan	1	1.35
Vietnam	1	1.35
Poland	1	1.35
Kosovo	1	1.35
Bosnia	1	1.35
Bulgaria	1	1.35

.

The questionnaire asked students to state which languages they use with different members of their family, which languages they use in their free time, and whether they have someone in their life they speak English with. Over the four classes participating in this study a diverse linguistic background and high degree of multilingualism can be observed. A total of 26 different languages are spoken by the students. As can be seen in Table 4, almost every student uses more than one language at home. Only two participants exclusively speak one language. When looking at predominantly used languages at home, surprisingly, languages other than German dominate participants' linguistic experience at home. 29.73 % of students primarily use Turkish with their family, followed by 20.27% using German or Serbian respectively. If all other languages are combined, even though individually they are only spoken by a few sometimes even just one learner, they make up the same percentage as Turkish (29.73%). Therefore, languages other than German constitute 79.73% of predominant languages spoken at home. This does not necessarily mean that these students never use German at home. 57 students of those 79.73% state that they still use German to some degree and only 7 students report that they never speak German at home. Yet, as German is the main language spoken in Austria, the reduced role of German in the homes of learners is still surprising.

Table 4: Languages used at home

Number of languages used at home	N	%
1	2	2.70
2	46	62.16
3	25	33.78
4	0	0.00
5	1	1.35
Predominant language used at home	N	%
Turkish	22	29.73
German	15	20.27
Serbian	15	20.27
other languages	22	29.73
Chechen	4	5.41
Kurdish	3	4.05
Arabic	2	2.70
Persian	2	2.70
Yoruba	1	1.35
Filipino	1	1.35
Croatian	1	1.35
Bosnian	1	1.35
Macedonian	1	1.35
Albanian	1	1.35
Bulgarian	1	1.35
Polish	1	1.35
Vietnamese	1	1.35
Chinese	1	1.35
Dari	1	1.35
German use at home	N	%
German only language used	2	2.70
German predominantly used	13	17.57
other language predominantly used	52	70.27
German not used	7	9.46
English use at home	N	%
yes	18	24.32
no	56	75.68

When it comes to English 24.32% of learners report using English at home and mostly with their peers, i.e. friends or siblings. Only 3 learners also speak English with either their mother or father. For one of them, English plays a significant role at home where it is used as an additional language to her first language with every family member. Even though such substantial English use differentiates her from other participants, she is still included in the study, because she was born in Austria and not an English-speaking country, the predominant language used at home is Yoruba[4] and not English, and she is not an outlier in terms of language level as she belongs to the 'Standard' group in English.

[4] Yoruba is a language spoken in West Africa and one of four official languages in Nigeria.

Though German does not feature significantly in most students' homes, the languages used in their leisure time, i.e. outside their home, present a different picture. Table 5 depicts the four languages learners reported using the most in their free time. As opposed to language use at home, every student speaks German to a varying degree in their leisure time with friends or during activities and for 27 teens German is the number one language here. English is the second most mentioned language, with 54.05% of participants using it in their free time. Strikingly, a higher percentage of students uses Turkish and Serbian here than at home. These results suggest that in this linguistically diverse group of learners, teens are more open to incorporate foreign languages their friends speak when talking to them during leisure time. Furthermore, especially German and English seems to be more frequently employed as a language for leisure time activities with peers than a home language.

Table 5: Most used languages in students' leisure time

	1st	2nd	3rd	4th	N	%
German	27	44	2	1	74	100.00
English	1	8	23	8	40	54.05
Turkish	17	6	1	3	27	36.49
Serbian	13	6	2	1	22	29.73

Overall, the group of participants in this study is culturally and linguistically very diverse. Almost all students are multilingual and languages other than German dominate their linguistic background at home. Even though German as well as English play a diminished role in the home of students, they are used by a majority of learners with peers as a language for leisure time activities.

6.1.2 Access to media

Whether engagement with EE is even possible in a digitalized world is highly dependent on access to the internet and digital devices. Therefore, the questionnaire inquired about the extent to which students have access to different types of media that allow them to consume and participate in digital EE content.

Table 6 reveals that all students not only have access to the internet, but they also possess different digital devices. 93.24% of students have an internet connection at any time at home and the 5 remaining students are sometimes able to access the internet from their home. At school there is a free WIFI connection and students who lack an internet connection at home can still engage in EE online in their leisure times, e.g. during breaks.

Table 6: Access to media

internet	N	%
yes	69	93.24
sometimes	5	6.76
no	0	0.00
phone	74	100.00
TV	74	100.00
PC or laptop	74	100.00
video game console	55	74.32
tablet	49	66.22
radio or stereo	26	35.14
e-reader	6	8.11

The participants in this study are well equipped to access digital media. Every student possesses or has access to a phone and a TV. Further, in part thanks to the initiative by the ministry of education to equip students with digital devices, every student in the study owns a laptop. These results show that especially a phone, TV, and PC or laptop seems to be standard equipment for students this age. More than half of learners also have a video game console (74.32%) and a tablet (66.22%). Fewer participants have access to a radio/stereo (35.14%), which is not surprising given that such devices are rather outdated. An even smaller number of students (6) owns an e-reader. But overall, it could be argued that the digital equipment is surprisingly extensive for children in this age-range and especially for a school in the 10th district of Vienna with a high proportion of students coming from low-income families.

6.1.3 General leisure time activities

One section of the questionnaire also focused on participants' general leisure time activities. What they do in their free time will certainly influence what their EE engagement looks like. Therefore, the trends observed in this section may hint at what EE activities are more and less popular amongst students.

Table 7 provides a breakdown of the reported frequency of general leisure time activities sorted by the category '(almost) daily'. Overall, over half of students engage in 8 of 13 activities on a weekly basis. Three activities are done by at least 50% of participants (almost) daily: using social media, watching videos on the internet, and listening to music. The fact that the top two activities take place online, points to the significance of digital and online spaces in students' leisure time. It is somewhat surprising that social media are used so regularly by the learners, given that they have not yet reached the required age to use apps such as Instagram, TikTok, or Facebook. With regards to previous studies on EE, it was expected that gaming would be more popular. Even though 60.91% of

participants reported 'playing games on your own' and 52.70% 'playing games with others' on a weekly basis, gaming was anticipated as one of the most favoured activities in this age group.

Table 7: Frequency of general leisure time activities

	(almost) never %	a few times a year %	a few times a month %	a few times a week %	(almost) daily %	Sum weekly %[5]	N
using social media (e.g. TikTok, Instagram, Facebook)	8.11	0.00	5.41	13.51	72.97	86.48	74
watching videos on the internet	1.35	4.05	8.11	20.27	66.21	86.48	74
listening to music	8.11	2.70	12.16	27.03	50.00	77.03	74
watching movies or series	8.11	4.05	27.03	24.32	36.49	60.81	74
playing video games on your own (e.g. console, laptop, PC,…)	9.46	4.05	24.32	24.32	36.49	60.81	73
watching TV	5.41	5.41	21.62	32.43	35.41	67.57	74
meeting friends	8.11	2.70	18.92	37.84	32.43	70.27	74
playing video games with others (e.g. multiplayer online games)	13.51	6.76	27.03	24.32	28.38	52.70	74
making music (singing or playing an instrument)	50.00	10.81	4.05	16.22	16.22	32.44	72
reading books or e-books	54.05	16.22	13.51	9.46	6.76	16.22	74
reading newspapers or magazines (also online)	58.11	13.51	16.22	6.76	5.41	12.17	74
going to the cinema	24.32	40.54	25.68	5.41	2.70	8.11	73
listening to audiobooks	72.97	17.57	8.11	0.00	1.35	1.35	74

There is a clear delineation between popular and unpopular general leisure time activities. Here, less frequent activities are characterized as such, if more than half of participants are never or only a few times a year occupied with them. These are: making music, reading books or e-books, reading newspapers or magazines, going to the cinema, listening to audiobooks. What stands out here is the infrequent engagement with reading. Previous studies on EE report that younger students tend to read less than older students (Avello et al. 2019: 84; de Wilde, Brysbeart & Eyckmans 2020: 177; Sundqvist 2009: 117; Warnby 2022: 131), and in this study this can also be observed when it comes to general leisure time activities. Whether this is true for EE for the present data sample is discussed below.

The Mann-Whitney u-test reveals differences in engagement between genders. Girls listen to music ($p = .002$, $d = .382$ [.136, .583]), make music ($p = .036$, $d = .267$ [.005, .495], and use social media ($p = .015$, $d = .257$ [-.002, .484] more frequently than boys. By contrast, male participants more often play video games with others ($p = .036$, $d = -.276$ [-.499, -.018][6]. Even though these effects are only small, the gender differences shown here are also reported by previous studies on EE (Hannibal Jensen 2017; Olsson and Slyvén 2015).

[5] The 'Sum weekly %' combines the frequencies of '(almost) daily' and 'a few times a week'.

[6] Note that the d-value shows the directionality of the effect with regards to female participants. A negative effect size points to a higher engagement of male participants than female participants.

This section provided additional information about the participants in this study, including their demographic and linguistic background, their access to media, and how they like to spend their leisure time in general.

6.2 Quantity and Quality of Extramural English engagement

In the following, the analysis shifts its focus to EE. How often students engage in EE and what activities are more and less frequently done will be illustrated here. Furthermore, differences between male and female participants will be examined. This section focuses on RQs 1-5 formulated above.

When it comes to EE activities participants are frequently occupied with, the data sample reveals a wide range in frequency of EE engagement. To attain an overall value for the frequency of EE engagement, an EE median score for each participant was calculated. The boxplot in Figure 2 shows the EE median score for the whole data sample (Mdn = 2). It displays that the middle frequency score of engagement is a few times a year across all activities and participants. Furthermore, the range in frequency of engagement covers all values from daily engagement (5) to no engagement (1). Over half of participants indicated very low overall EE median scores, with 35.14% of participants reporting the lowest EE median score of 1 (N = 26) and 17.57% an EE median score of 2 (N = 13). Overall, learners report 0 to 36 daily EE activities with a mean of 8.03 [9.80, 6,25]. 13 participants indicate that they do not engage in any daily EE pursuits in their leisure time.

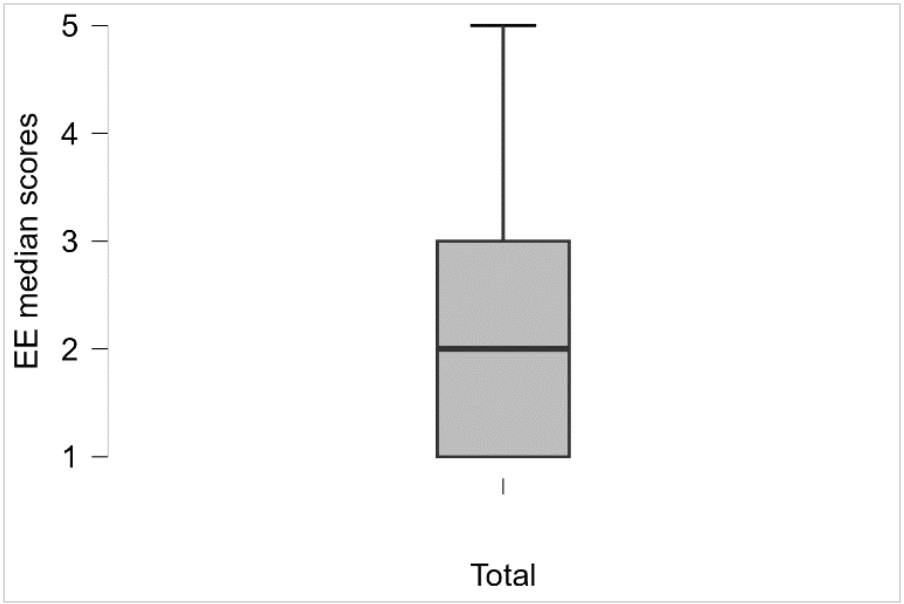

Figure 2: *Boxplot showing median and interquartile range for EE median scores.*

When looking at weekly EE engagement, which combines both daily and weekly activities, there are still three participants that do not do any of the activities mentioned in the questionnaire on a weekly basis. Strikingly, one of those students is the sole participant who indicated to never be engaged in EE, which is surprising given the presence of English in our daily lives. Such a result could be due to the participant not having any motivation to fill out the questionnaire and quickly ticking (almost) never for every activity. However, it can be assumed that this particular student has at least no real interest in EE and does not consciously seek opportunities to encounter English in their free time. Yet, the vast majority of students frequently engage in EE, with 82.43% doing at least one activity on a daily basis. An even higher percentage, namely 95.94%, are occupied with EE on a weekly basis; however, they do fewer activities on a weekly basis, as the mean of weekly activities is 6.57 [7.71, 5.42] compared to 8.03 daily activities.

When comparing these results to Schwarz' study (2020) on older students, participants in this study engage in EE to a lesser extent than their older counterparts. The mean of daily and weekly activities as well as the number of learners never engaging in such activities supports the hypothesis that younger learners are less frequently involved with EE than older students.

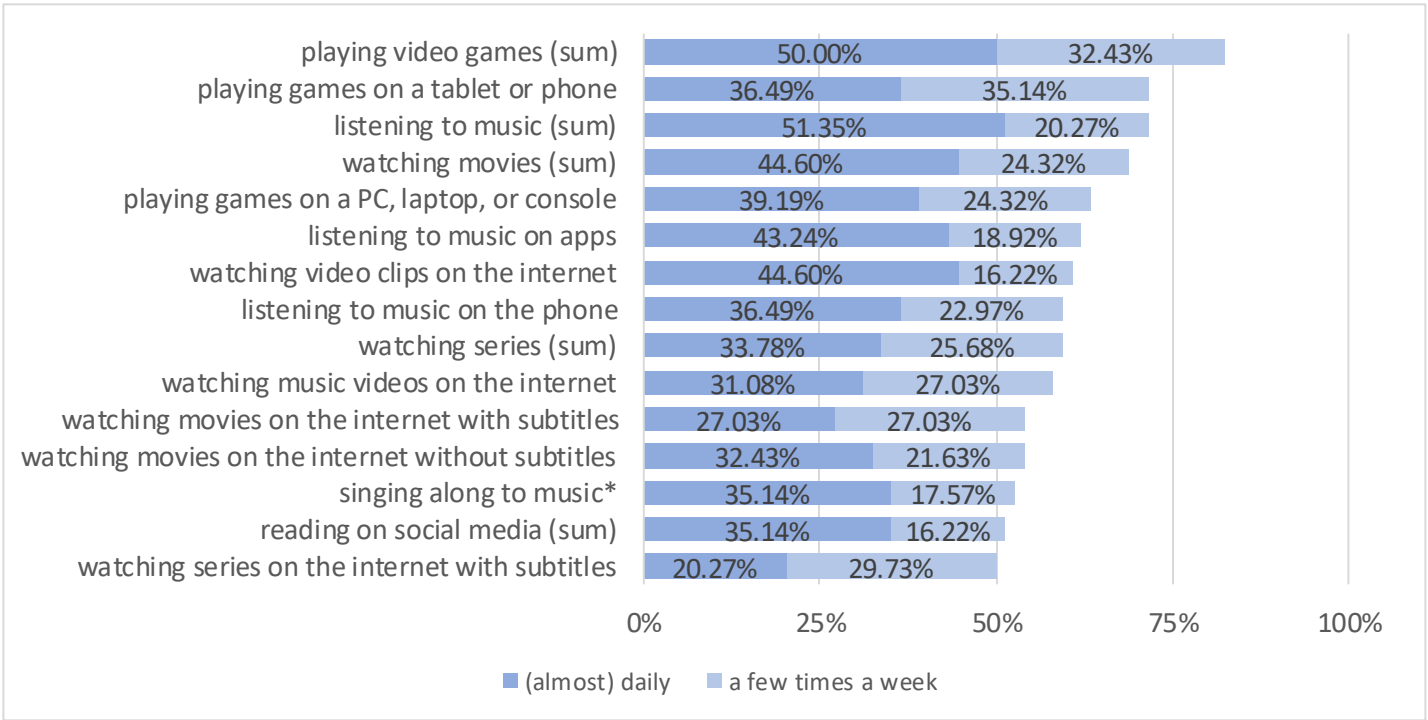

Figure 3: *Most frequent EE activities at least 50% of participants (N=74) engage in on a weekly basis.*

In terms of quality of EE, the results show that gaming, listening to music, and watching audiovisual input are most popular with learners in this study. The most frequent activities listed in Figure 3 are those in which over half of the participants engage in at least once a week. The label 'sum' next to an activity indicates a summary variable that combines multiple questionnaire items. Gaming is the

most favoured activity, with 72.43% playing video games overall at least once a week. Gaming also seems to be popular independent of which device is used, as playing games on a tablet or phone and playing games on a PC, laptop, or console are featured in the top five EE activities. However, one item in the questionnaire also inquired about playing multiplayer online games, and this type of gaming is not included here, as only 45.94% of participants engage in it on a weekly basis.

Next, 71.62% of participants listen to English-language music on a weekly basis. This is slightly less than students reported for listening to music as a general leisure time activity (77.03%), which is to be expected as leaners certainly also listen to music in their first or other languages. However, the majority of consumed music seems to be in English. When only looking at daily engagement, listening to music is even more popular than gaming, with 51.35% of participants being occupied daily with English-language music. Students listen to music in many ways, on apps, their phone, or via music videos on the internet. Another activity connected to music listed in Figure 3 is singing along to music, with 52.71% frequently engaging in this activity. Singing along to music represents the only activity here where students assuredly use English productively (indicated in Figure 3 with *).

The third most frequent EE activity is watching movies, with 68.92% watching them on a weekly basis. The exact same percentage of students (54.06%) watch movies on the internet with or without subtitles, which means that some students indicated doing both, not preferring one or the other. Previous studies on EE revealed that younger students seem to prefer watching movies, series, or TV with subtitles, but that is not the case in this study as students do not favour one over the other (Bollansée & Puimège 2020: 206, de Wilde, Brysbeart & Eyckmans 2020: 177, Tam & Reynold 2022: 21). In general, watching audiovisual input is featured heavily here, as watching videoclips, series, and music videos are also among the most frequent EE activities.

Lastly, reading on social media is done weekly by 51.36% of participants. That includes reading direct messages, comments, and status updates. Given that using social media is the number one general leisure time activity of participants, it is interesting that this is the only activity featured amongst the most popular EE activities that takes place on social media. However, it can be argued that watching video clips on the internet can very well also take place on social media platforms, and that students interact less with written and more with audiovisual content on social media platforms.

Taking a closer look at the frequency of engagement in specific EE activities, this section will focus on activities that most students encounter at least on a monthly basis. First of all, when it comes to

monthly engagement, students use language more productively. A total of 5 productive activities are presented in Figure 4. The number of productive EE activities increases with decreasing frequency of engagement, which points to the fact that learners prefer receptive over productive EE activities.

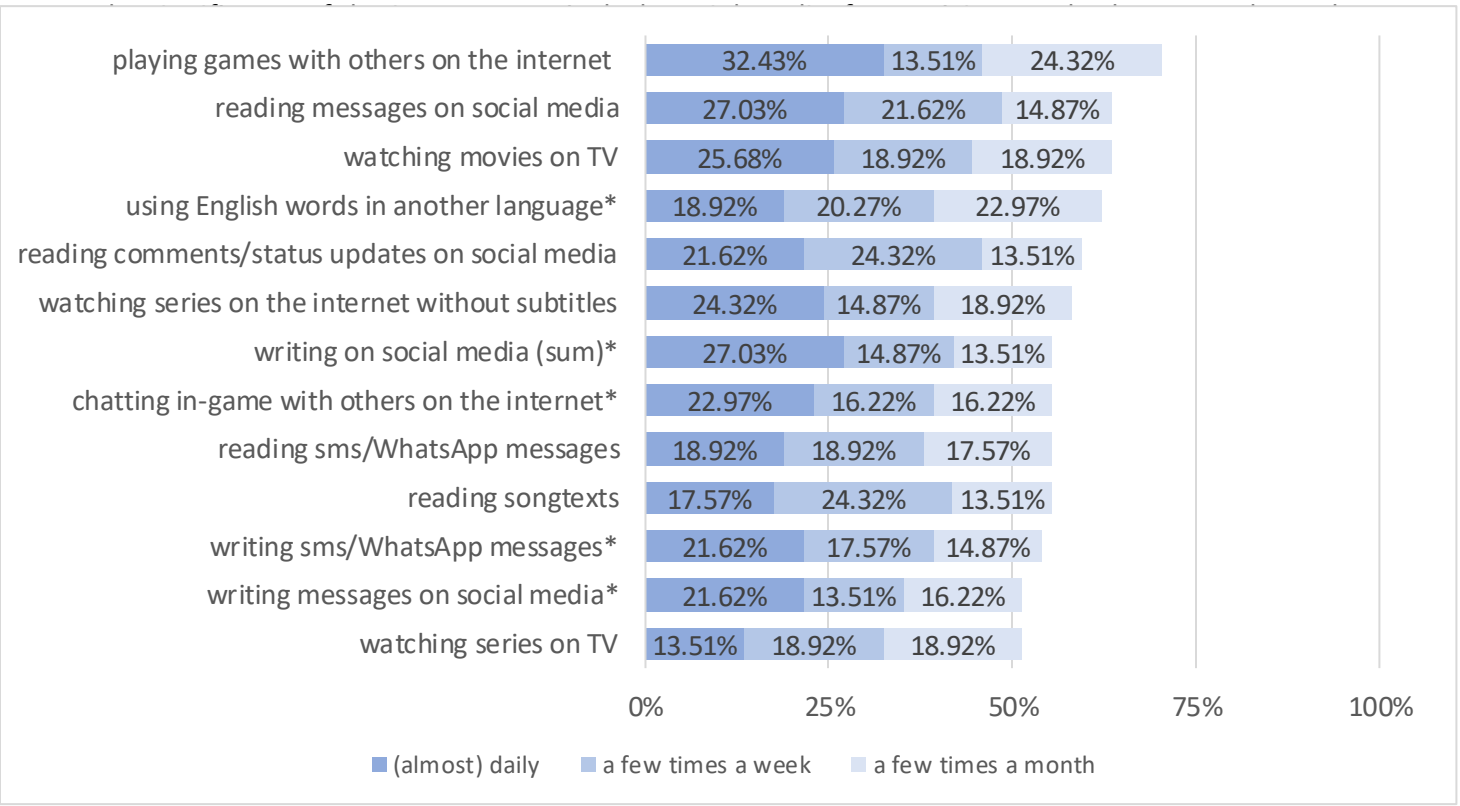

Figure 4: *EE activities at least 50% of participants (N=74) engage in on a monthly basis.*

Another relatively popular use of social media or the internet is communicating through direct messages, comments, or status updates. 55.38% of participants report to write at least monthly on social media. Especially popular seems to be direct messages, with 51.35% of learners indicating that they write messages on social media at least once a month. Another key form of communication is through SMS or WhatsApp, with 54.06% of students sending such messages monthly. Give the importance of WhatsApp in daily life, students certainly use WhatsApp overall to a higher frequency, but most likely use languages other than English. As with EE reading, EE writing is for most students in this study restricted to short quick messages or other texts written online. Consistent with the trend of frequent EE activities happening online, watching movies or series on TV is less popular than watching movies or series on the internet. 44.60% watch movies and 32.43% watch series on TV on a weekly basis, compared to 54.06% watching movies with or without subtitles on the internet and 50% watching series with subtitles on the internet.

Following this analysis of the popular, more frequent EE activities, I will illustrate which EE activities students engage in less frequently. Figure 5 displays all activities which more than half of students only do a few times a year or never. Several trends can be observed here. First, productive language use such as writing or speaking is rarely done as an EE activity. The majority of students rarely or never write, for example, in a diary (91.90%), e-mails (74.33%), or texts such as short stories, manuals, or articles (64.87%). Like EE reading, writing as an EE activity is restricted to short texts that are mostly written on social media or as messages, and it rarely involves longer texts. As to speaking, contrary to the trend towards online activities observed above, fewer learner rarely engage in face interaction (52.71%) rather than on the phone (60.81%) or online via videochat (68.92%).

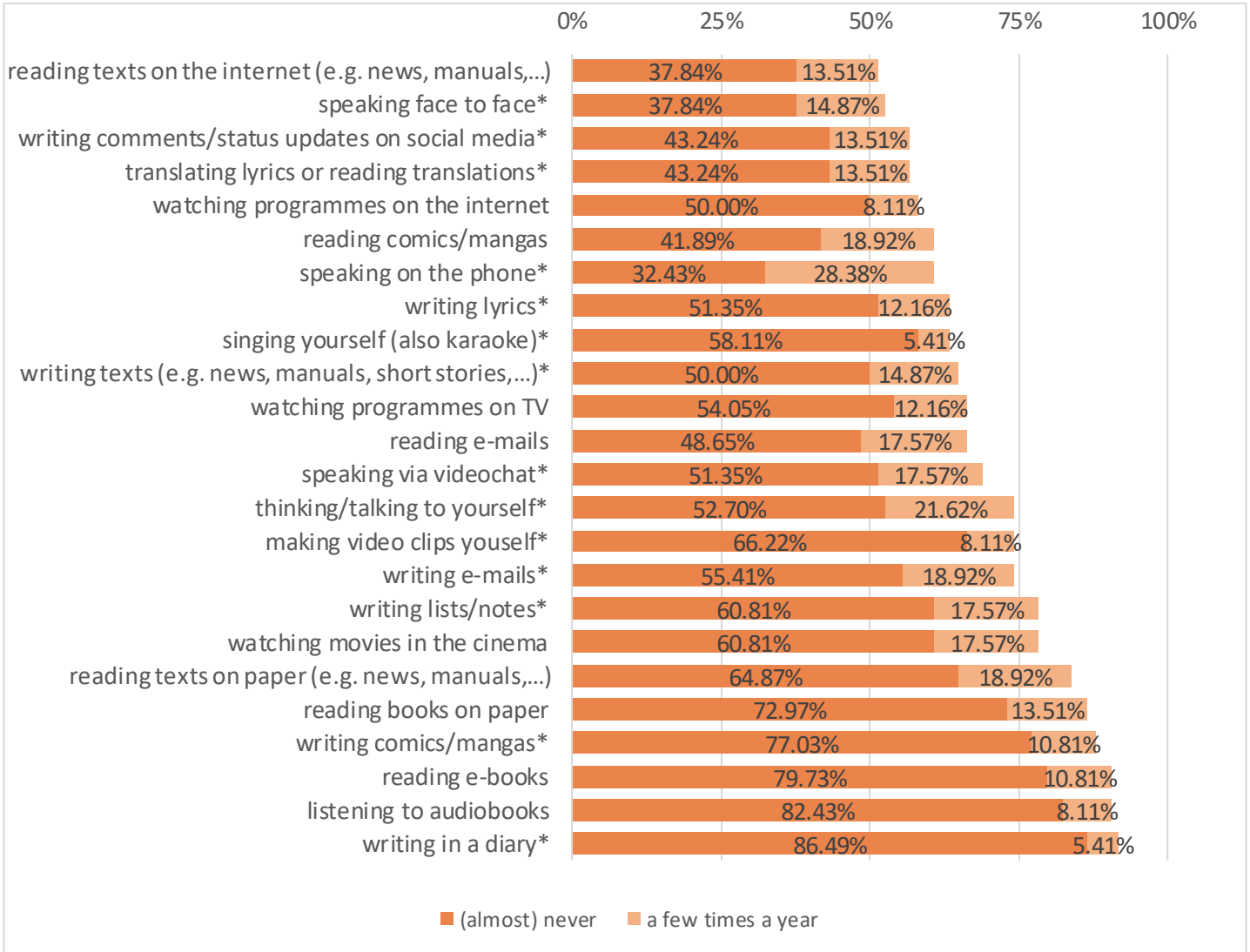

Figure 5: *Least frequent EE activities over 50% of participants (N=74) engage in only a few times a year or (almost) never.*

Next, EE reading other than short texts on social media plays a diminished role in students leisure time. Especially books seem to not be popular with the participants in this study, with over 70% never reading English books or e-books in their leisure time. In addition, listening to audiobooks is the second to last least frequent activity. Other genres are also featured in the list of least frequent

reading activities. Over 60% never or only a few times a year read comics/mangas or emails and over 80% rarely read texts such as the news, manuals, or other text types on paper. Just over half of participants (51.35%) seldom read texts (e.g. news, manuals, etc.) on the internet, which leaves 48.65% who engage in this activity regularly. This again points to the significance of the internet in how students spend their free time.

Lastly, offline activities are less frequently carried out by students than activities that take place online. This was already pointed out with reading and writing activities, but it also becomes apparent when looking at what participants watch in their free time. Activities that do not involve the internet, such as watching programmes on TV or watching movies in the cinema are among the least frequent EE activities. This could be due to the prevalence of on-demand streaming via the internet. Learners seems to prefer choosing what they want to watch at any time over predetermined series, movies, or programmes on TV or in the cinema.

Even though these EE activities are less popular, some students do engage in them quite regularly. Such activities could be termed as 'niche' activities, as they are not the mainstream within this particular group of learners, and they potentially hint at a student's higher language level as they might require higher level language processing. Looking at the five least frequent activities, in which over 80% of participants rarely or never engage in, reveals that a total of 12 students do at least one of those activities on a weekly basis. Interestingly, all of them are female. In terms of writing, two girls keep an English diary and write in it every day and one student writes comics or mangas on a daily basis. There are no learners who listen to audiobooks regularly, only some do it a few times a month. When it comes to reading, e-books and books are read daily by two girls respectively, and another pair of participants reads books on a weekly basis. The two learners who often read books seems to be avid readers of other genres as well, as they are also engaged in reading texts such as the news, manuals, and others on paper on a daily basis. Overall, four participants engage in two of the five least frequent activities in a regular basis and one student even does three of them frequently.

Learners who engage in 'niche' activities from the list in Figure 5 can also be divided into groups. A total of four different groups were created from activities that might yield insightful results in terms of their correlation to LD scores and difference to other participants. These are: the avid readers, the writers, the talkers, and group self-talk. All groups include participants who engage in certain activities on a daily or weekly basis. The group of 'avid readers' only includes those students who frequently reading books or e-books. The reading of shorter texts such as the news or others is not

included here, because the focus of this group is on extensive reading. Next, 'the writers' are students who engage in writing of comics/mangas, texts such as short stories or others, e-mails, lists/notes, and keeping a diary. Both of these groups do not include reading or writing on social media, as the majority of students is involved with such activities on a regular basis. All speaking activities that involve other people, i.e. speaking face to face, on the phone, and via videochat, are included in 'the talkers' group. Lastly, the group 'self-talk' is made up of those participants who frequently think or talk to themselves in English. Self-talk or an inner monologue can be in indicator of cognitive L2 fluency and could point to an advanced immersion in a L2, this is why this group is included here (Segalowitz 2016). These groups as well as individual learners who engage in 'niche' activities are going to be revisited when looking at correlations between certain EE and LD scores and differences between certain groups of learners.

To conclude the investigation of the quality of students' EE engagement, it can be said that the results support findings of previous studies. First, learners engage more in receptive than productive EE activities. Second, when comparing the results to the study by Schwarz (2020), the younger learners in this study reported to play video games on a more frequent basis than older learners in her study. While previous studies reported that younger learners make more use of subtitles when watching movies or series, the results here do not support these findings. Furthermore, the least frequent EE activities in this study correspond to findings of previous studies, namely that students rarely engage in EE writing and reading as well as offline activities.

6.2.1 Gender differences

The are some differences between female and male participants in terms of quantity of EE engagement. First, the boxplots in Figure 6 reveal that females exhibit a higher EE median score ($Mdn_{female} = 3$, $Mdn_{male} = 2$), suggesting that girls are more frequently engaged in EE than boys ($p = .055$, $d = .251$ [-.009, .479]). Even though the d-value indicates a small effect size for the difference in EE median scores, the results are statistically not significant and the wide 95% CI points to some uncertainty regarding the magnitude of gender differences in the frequency of EE engagement. However, all participants with an EE median score of 5 are female, whereas the highest median score of male participants is 4. This shows that the range in quantity of EE is higher with female participants and those participants who have the highest frequency of EE engagement are all girls as well. Yet, when looking at the number of daily, weekly, monthly, yearly, and never engaged in activities, the Mann-Whitney u-test indicated no significant differences between the two groups. Girls have a slightly higher mean of daily and weekly activities, while boys display a marginally higher mean of

monthly, yearly, and never engaged in activities. To conclude, while girls and boys do not differ to a statistically significant extent in terms of quantity of EE engagement, female participants have a higher EE median score, and it can be assumed that they differ from male learners at least to a small effect.

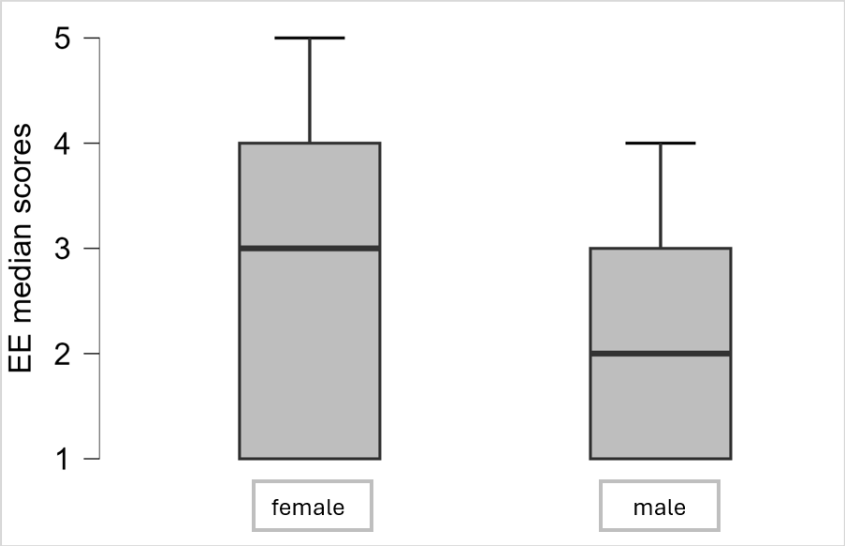

Figure 6: *Boxplot showing median and interquartile range for EE median scores according to gender.*

When it comes to the particular activities students do frequently or infrequently, the analysis shows that girls and boys differ slightly in terms of quality of EE. Overall, females engage in more EE activities on a weekly basis[7]. The top five most frequent activities shown in Figure 7 are similar for boys and girls with some differences. It can be said that the three summary variables in Figure 7 are the most popular EE activities of female participants, with 80.55% listening to music, 77.78% playing video games, and 69.44% watching series weekly. The same is true for boys; however, first of all, the order of most popular activities differs and second, boys prefer watching movies over watching series. Playing video games is the number one activity (86.85%), which is most often done on a PC, laptop, or console, followed by listening to music (68.43%), and watching movies (68.42%). The main take away is that, while girls listen to music, especially on the phone, more regularly ($p = .006$, $d = .361$ [.113, .567]), male participants play video games more often, especially on a PC, laptop, or console ($p = .006$, $d = -.356$ [-.563, -.107]).

[7] As described above, those are activities that at least 50% of participants engage in on a weekly basis.

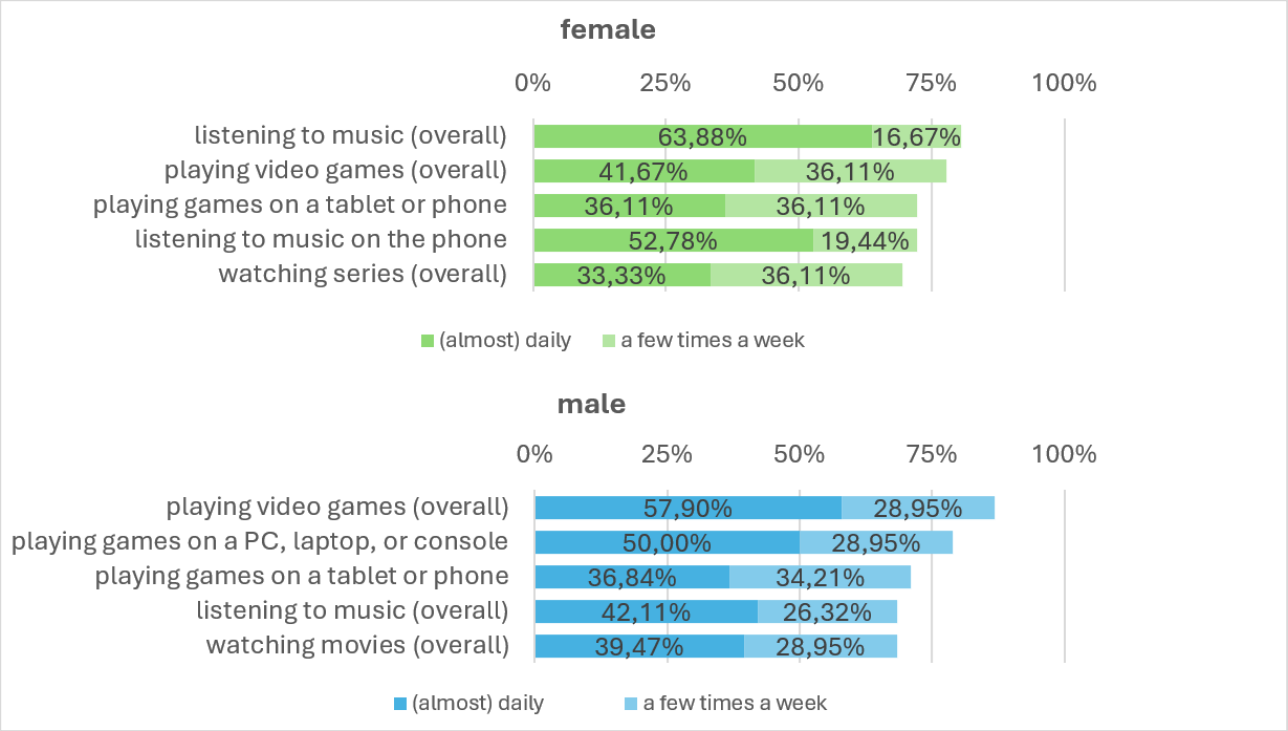

Figure 7: *Top five weekly EE activities according to gender.*

Even though the most frequent EE activities do not differ much, there are some EE free time pursuits where the Mann-Whitney u-test shows a marked difference in terms of engagement between boys and girls. Apart from playing video games on a PC, laptop, or console, all other activities that show statistically significant discrepancies between genders are done more frequently by girls. The most notable difference in terms of effect size can be observed in connection to WhatsApp messages or sms, where reading of such messages shows a small to medium and writing a medium effect size. Girls read more English WhatsApp messages or sms ($p = .005$, $d = .378$ [.131, .582]) and engage in writing of such messages to a considerably greater extent ($p < .001$, $d = .524$ [.305, .690]). There are also other activities that show a statistically significant higher engagement of female participants, but to a smaller extent in terms of effect size: using English words in another language ($p = .022$, $d = .303$ [.048, .521]), translating or reading translations of English lyrics ($p = .026$, $d = .287$ [.030, .508]), and writing in a diary ($p = .005$, $d = .228$ [-.033, .460]). Furthermore, while the p-value shows no statistical significance, small effect sizes can be observed for writing lists or notes ($p = .053$, $d = .230$ [-.030, .462]), speaking on the phone ($p = .052$, $d = .254$ [-.005, .482]), and speaking face to face ($p = .053$, $d = .255$ [-.006, .483]).

With regards to the hypotheses formulated above, these results suggest that boys do play more video games, whereas girls are more involved with English music through listening and engaging with lyrics. In addition, female learners are more active when it comes to using English to

communicate via WhatsApp or sms, and to write down everyday occurrences and things to remember in a diary or on lists/notes. There also seems to be a difference in speaking activities between boys and girls, albeit to a lesser extent. Previous studies also suggested that boys report a higher quantity of EE. However, Cohen's d suggests a higher frequency of EE engagement for female participants compared to boys to a small effect. Yet, the p-value does not reach statistical significance. Therefore, in contrast to previous studies, it can be assumed that there is no marked difference between genders in terms of quantity of EE.

6.3 Extramural English and lexical diversity

The following section aims at answering RQs 6-9 as well as the overarching research question of this thesis. First, the LD scores for each text, the mean of both texts' scores, and potential differences between groups of students will be presented. Next, I will go into detail on possible correlations between LD scores and different EE variables. Lastly, students grouped according to EE engagement are going to be compared with regards to their LD scores.

6.3.1 Lexical diversity scores

The participants texts exhibit LD scores ranging from .393 to .826 with mean MATTR scores that are considerably lower than in previous studies. Figure 8 shows boxplots of MATTR scores for the mean of both types as well as the mean for each text separately. The median scores for the mean of both texts is .610 with scores ranging from .411 to .800 and a mean (M) of .605. This wide variety in LD illustrates the heterogeneity of participants' language level. When comparing the two different text types, the descriptive texts about students' favourite sport exhibit a slightly lower median of .599 (M = .594), while learners scored a higher median of .627 (M = .615) in the narrative about an experience in bad weather. This might suggest that the genre and nature of the writing task influences students' writing performance in terms of LD. Additionally, these results might reflect a tendency of sports-related vocabulary being more constrained, thus limiting the range of lexical items that could potentially be used. Narrative writing might offer more opportunities to use a more diverse vocabulary as learners recount personal experiences and emotions.

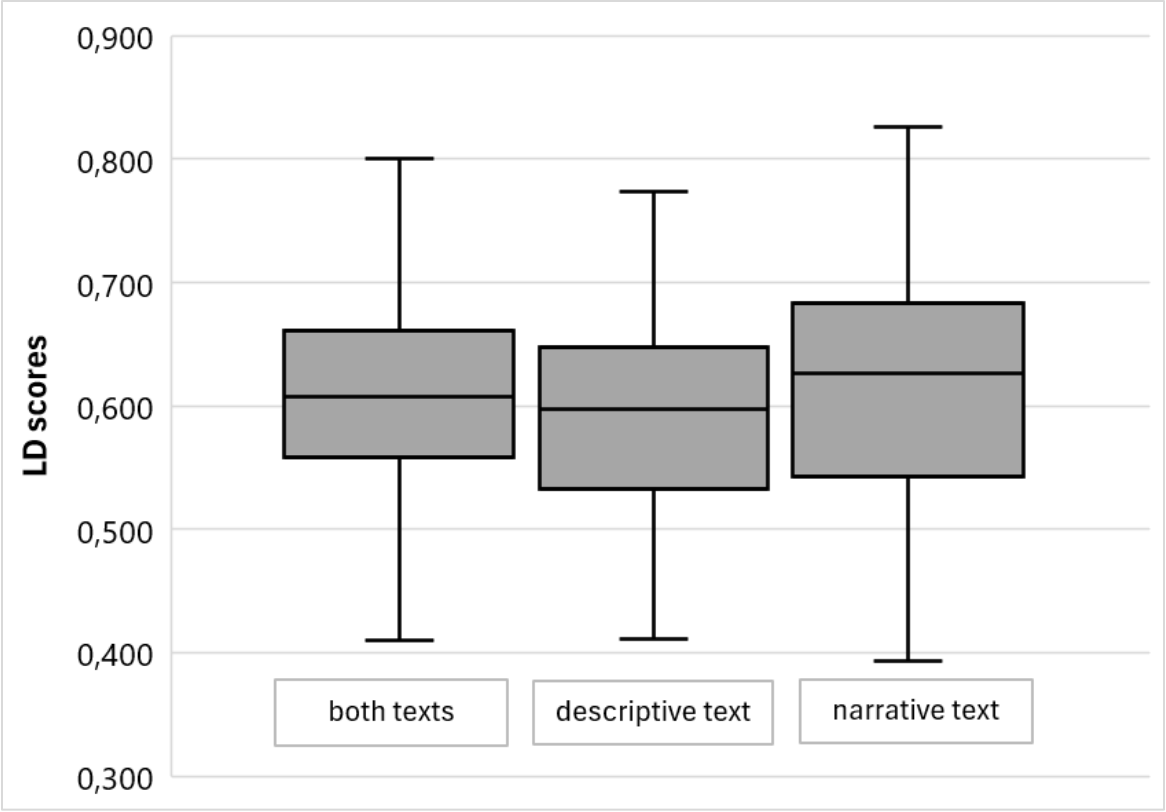

Figure 8: *Boxplot showing median and interquartile range for LD scores of both texts, descriptive text, and narrative text.*

In comparison to the findings from prior studies, the present study's lower mean MATTR scores align with differences in the age and proficiency level of participants. For example, previous studies by Kyle, Crossley and Jarvis (2021) and Zenker and Kyle (2021) found higher MATTR mean scores of .740 and .730 respectively in the writing of learners of English. These studies, however, involved older university or upper-secondary students, who are expected to possess more advanced language skills compared to participants in this study. Moreover, the genre of writing could also play a significant role in higher MATTR scores. Both studies investigated the LD of argumentative texts. This genre requires writers to argue a viewpoint, present evidence, elaborate on abstract concepts, and engage in critical thinking, which might offer more opportunities to use a lexically diverse vocabulary than narrative or descriptive texts, which may not necessarily demand the same degree of LD. Even though no study examined LD scores across genres, the higher MATTR mean might thus be influenced by the text genre.

When looking at different groups of students and how they differ in terms of MATTR scores, higher level students and female participants produced text with a higher LD. Students in a 'Mittelschule' are divided into two groups, with higher level, or more correctly higher achieving, students belonging to the 'AHS' group and the others to the 'Standard' group. The texts of 'AHS' learners show a significantly higher MATTR mean score than those of 'Standard' learners. This is especially salient

when examining the results of the Mann-Whitney u-test illustrated in Table 8, which show a statistically significant difference in LD scores between the two groups with large effect sizes for the LD mean scores of both texts and the descriptive texts. The 95% CI for these variables are rather narrow, which makes the assumption of larger effect sizes more reliable. In addition, LD scores of 'AHS' and 'Standard' students' narrative texts differ to a medium effect.

Table 8: Differences in LD scores between 'AHS' and 'Standard' students

	p-value	Cohen's d	95% CI
both texts	p < .001	.855	[.749, .918]
descriptive text	p < .001	.817	[.689, .896]
narrative text	p < .001	.691	[.498, .819]

The analysis of MATTR scores reveals that although female participants tend to have slightly higher scores than male participants across all texts, these differences are not statistically significant. Table 9 illustrates that the mean MATTR score of female participants is higher in the descriptive text (M_{female} = .621, M_{male} = .577), the narrative text (M_{female} = .629, M_{male} = .602), and when examining the mean scores of both texts (M_{female} = .621, M_{male} = .590). However, despite these numerical differences, these variations are not significant enough to suggest a meaningful difference in LD between male and female learners. This is further supported by the small effect sizes observed in the data, which suggest that the impact of gender on MATTR scores is minimal.

Table 9: Differences in LD scores between female and male participants

	p-value	Cohen's d	95% CI
both texts	p = .096	.226	[-.035, .458]
descriptive text	p = .069	.246	[-.013, .475]
narrative text	p = .272	.149	[-.114, .393]

6.3.2 Correlation between lexical diversity scores and Extramural English engagement

After describing the LD scores of students' texts, this section deals with possible correlations between lexical diversity and EE. This represents the core focus of this thesis. This thesis uses Kendall's tau (τ) as a correlation coefficient. In a first step, to investigates to what extend quantity and quality of EE influences LD, EE median scores, summary variables, and engagement in specific EE activities will be discussed with regards to correlations with LD. Next, the discussion shifts to

differences in LD between students with low and high EE engagement and how learners belonging to 'niche' groups differ from the rest in terms of LD.

6.3.2.1 Quantity of EE vs. quality of EE

To get an overall indication of how the quantity of EE correlates with LD, the relationship between EE median scores, as a global value of frequency of EE engagement, and MATTR scores is examined. The correlation analysis displayed in Table 10 reveals that EE median scores can be an indicator of LD. It shows a statistically significant positive correlation between quantity of EE and LD mean scores of both texts with a small to medium effect (τ = .333, [.193, .473], p < .001). Further analysis shows that this positive relationship extends to both types of texts. MATTR scores of the descriptive text (τ = .284, [.141, .427], p < .001) and narrative text (τ = .302, [.156, .449], p = .001) also show positive correlations, albeit to a smaller effect than the LD mean of both texts. These results suggest that with increasing EE engagement students LD in writing, expressed through MATTR scores, also increases. This supports the hypothesis of RQ 7, namely that the quantity of EE engagement positively correlates with LD in writing. While the effect sizes are not large, they are nonetheless significant, indicating that how frequently students engage in EE plays a role in enhancing their LD in writing.

Table 10: Correlation between EE median scores and LD scores

	both texts	descriptive text	narrative text
EE median score	τ = .333	τ = .284	τ = .302
	[.193, .473]	[.141, .427]	[.156, .449]
	p < .001	p < .001	p = .001

To investigate what role the quality of EE plays in connection with LD in writing, potential correlations between the six summary variables and LD scores are going to be examined. These summary variables include most of the most frequent EE activities. Furthermore, correlation analysis between all EE activities and LD scores is going to be conducted to investigate whether certain activities trump quantity of EE engagement in terms of effect size.

Table 11 illustrates that the correlation analysis shows only small effect sizes and the relationship between only two variables and all LD scores reaches statistical significance. Reading on social media seems to have the most pronounced effects. The correlation coefficient suggest a small to medium correlation with MATTR scores for the mean of both texts (τ = .330, [.186, .470], p < .001), the descriptive text (τ = .274, [.109, .426], p = .002), and the narrative text (τ = .306, [.149, .467], p < .001). Next, writing on social media correlates with LD scores to a small effect. The greatest effect

size can be observed in connection to the narrative text (τ = .266, [.108, .422], p = .002). Listening to music also positively correlates with the mean LD of both texts (τ = .227, [.056, .392], p = .011) and the narrative text (τ = .241, [.064, .413], p = .007). No statistically significant correlation can be found in connection with the descriptive text. The same is true for watching movies. For playing video games and watching series, the relationship with LD scores does not show statistical significance and effect sizes are so small that no positive correlation can be assumed.

Table 11: Correlation between summary variables and LD scores

	both texts	descriptive text	narrative text
Listening to music	τ = .227	τ = .159	τ = .241
	[.056, .392]	[-.008, .320]	[.064, .413]
	p = .011	p = .075	p = .007
Watching movies	τ = .197	τ = .140	τ = .205
	[.056, .332]	[-.017, .296]	[.054, .360]
	p = .027	p = .115	p = .021
Watching series	τ = .143	τ = .101	τ = .153
	[-.004, .301]	[-.059, .272]	[-.006, .325]
	p = .103	p = .248	p = .081
Playing video games	τ = .140	τ = .127	τ = .132
	[-.037, .302]	[-.035, .303]	[-.057, .297]
	p = .123	p = .161	p = .144
Reading social media	τ = .330	τ = .274	τ = .306
	[.186, .470]	[.109, .426]	[.149, .467]
	p < .001	p = .002	p < .001
Writing social media	τ = .261	τ = .185	τ = .266
	[.085, .434]	[.003, .351]	[.108, .422]
	p = .003	p = .034	p = .002

The correlation coefficient across all variables shows smaller effect sizes for the descriptive text compared to the narrative one and the mean of both texts. This might suggest that these activities have a greater positive influence on students' LD when writing narrative texts rather than descriptive ones. In the analysis that follows, effect sizes are larger for the narrative text than the descriptive one and effect sizes for the mean LD score of both texts are usually in between. This is why the mean LD scores of both texts will be used to discuss the results of correlation analysis below, except when results for the narrative text and descriptive one exhibit significant differences from the norm.

When examining specific EE activities and their relationship to LD, the correlation analysis indicates that no individual activity demonstrates a markedly higher positive correlation than the EE median score. However, several activities show positive correlation with an effect size similar to the EE median score. When considering the correlation with the mean of LD scores of both texts, speaking to other people in person shows the highest effect size (τ = .379, [.219, .519], p < .001), followed by watching videos on the internet (τ = .345, [.184, .481], p < .001), listening to music on the phone (τ

= .333, [.176, .473], p < .001), reading comments or status updates on social media (τ = .323, [.179, .453], p < .001), and reading messages on social media (τ = .322, [.171, .458], p < .001). These activities appear to influence the LD of participants' texts to a similar effect than the EE median scores. All other EE activities included in the questionnaire display only small or negligible correlations with LD, and for some, the correlation analysis does not reach statistical significance.

A factor in the quality of EE is also the need for higher language processing, and activities that require such intensive language use are often productive in nature. This is why effect sizes of the correlation analysis of receptive/productive activities and LD scores were compared to discern whether productive language use has a markedly higher influence on LD. The results show that there is no discernible difference between the influence of receptive versus productive EE activities on LD scores. While four of the five activities that correlate with LD to an effect comparable to the EE median score are receptive in nature, a higher number of productive activities demonstrate a small correlation, whereas more receptive activities exhibit a correlation of less than small magnitude. Moreover, when considering the quality versus quantity of EE engagement, the correlation coefficient reveals that only a few specific activities display an effect size comparable to the median EE score. The majority of EE activities are, therefore, surpassed in impact by a high quantity of overall EE exposure.

The findings reported in connection with the influence of both the quantity and quality of EE on LD scores reveal similar effect sizes for EE median scores, as an indicator of quantity of EE, and some specific activities, such as reading on social media, speaking to other people in person, watching videos on the internet, and listening to music on the phone. The relationship between LD and other summary variables or specific activities show effect sizes smaller than the EE median scores, or the correlation analysis often does not reach statistical significance. This suggests that while certain EE activities may contribute to a higher LD, they are not consistently more impactful than the general frequency of EE engagement. Regarding the hypothesis about quantity versus quality of EE and its influence on LD formulated above, the data provides no statistical basis so far to suggest that quality of EE has a greater effect than frequency of engagement. Although activities of a higher quality, such as engaging in conversations, positively correlate with LD, the data does not support the assumption that these activities have a greater effect than simply engaging frequently with EE, regardless of the quality of engagement.

6.3.3 Differences between groups

To gain further insights into what EE practices have the highest impact on LD, students are put into different groups according to the frequency of EE engagement and their regular participation in niche activities. This approach may also help to determine whether quantity or quality of EE plays a more impactful role in participants' LD scores.

6.3.3.1 Low vs high EE

To compare participants according to the frequency of their EE engagement, groups based on EE median scores were created, using the overall average score of the sample (M = 2.365) as a reference point. The groups were categorized as follows:

- Low EE group (n = 39): This group consists of students with an EE median score below the average, i.e. those who indicated a median EE engagement of (almost) never or a few times a year
- Medium EE group (n = 20): Participants in this group have a median EE score of 3, indicating a moderate level of EE engagement around the average score.
- High EE group (n = 15): This groups includes participants with scores of 4 or 5. These students are the most engaged and reported a daily or weekly median EE engagement.

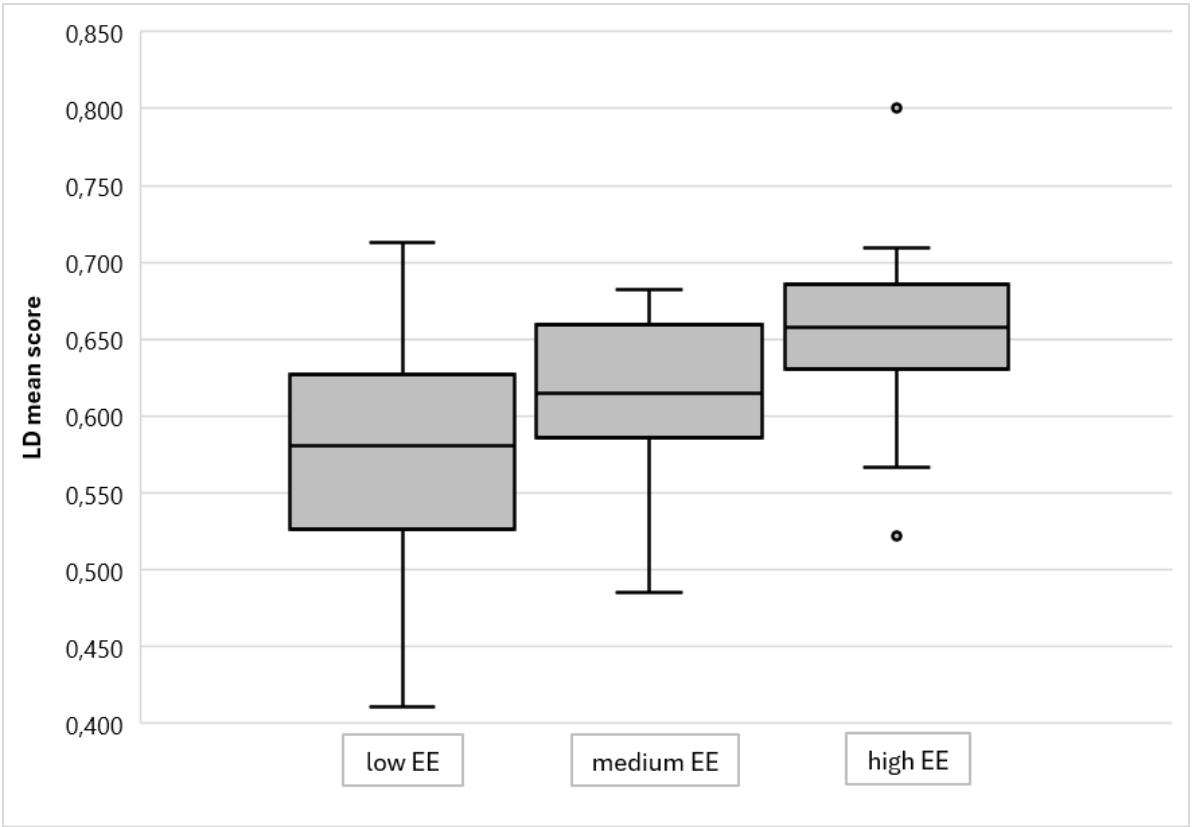

Figure 9: *Boxplot showing median and interquartile range for LD scores of participants grouped according to frequency of EE engagement*

Figure 9 illustrates the comparison between these groups in terms of their LD. The boxplots show the distribution of MATTR mean scores of both texts for each group. The low EE group exhibits the lowest median (Mdn_{low} = .581 [.555, .602]) and mean (M_{low} = .579), followed by the medium EE group (Mdn_{medium} = .615 [.594, .641], M_{medium} = .618). The high EE group has the highest LD mean scores (Mdn_{high} = .658 [.620, .690], M_{high} = .655).

These findings indicate a positive correlation between the frequency of EE engagement and LD, with the high EE group showing notably higher levels of LD in their texts. This aligns with the results of the correlation analysis between EE median scores and LD illustrated above.

Participants in the low EE group, however, display a wide range of MATTR scores, with some reaching scores higher than the median score of the high EE group. It can therefore be said that, even though they indicated to rarely engage in EE, some students were able to produce texts with a high LD. Yet, these results show that LD median scores are increasing with higher EE engagement. There are also two outliers indicated in the high EE boxplot in Figure 9. One of these is a participant who scored by far the highest MATTR score with .800. This particular learner will be revisited in the next section about niche groups.

Further analysis showed that these results are supported by the Mann-Whitney u-test, which reveals statistically significant differences between the low, medium, and high EE groups. Since the Mann-Whitney u-test can only be used to compare two groups, separate tests were conducted to assess the differences between each pair of groups, low and high EE group, the low and medium EE group, and the medium and high EE group. All groups differ from each other to a statistically significant extent, but to a varying effect size. The highest effect can be observed between the low and high EE group (p = .001, d = -.564 [-.284, -.756], indicating a robust difference in LD between students who rarely engage in EE and those who engage frequently. The analysis for low vs. medium EE group (p = .037, d = -.336 [-.036, -.581] and medium vs. high EE group (p = .049, d = -.397 [-.028, -.671] also showed statistically significant differences but only small effect sizes. These results suggest that while there are measurable differences in LD as EE engagement increases from low to medium and from medium to high, these differences are less pronounced compared to the substantial gap between the low and high EE groups.

6.3.3.2 Niche groups

The analysis of niche groups compared to the rest of participants, revealed that, while quantity of EE engagement does matter in terms of correlation between EE median scores and LD, specific

activities or patterns of EE engagement show a higher positive relationship with LD in writing. The groups under investigation here are made up of students who are frequently involved in activities the majority of participants reported to rarely do in their free time. To analyse differences between niche groups and the rest, a Mann-Whitney u-test was conducted. Results are presented in Table 12 below.

The avid readers differ from the rest to the highest magnitude, as measured by Cohen's d. The analysis shows a medium effect size (d = .693) in the upper range close to the threshold of larger effect sizes defined by Plonsky and Oswald (2014: 889). However, the 95% CI exhibits a very wide range, [.262, .894], where the true mean could potentially lie close to the lower boundary in the area of small effect sizes. This could be due to the small number of members in this group. Even though there is some uncertainty with regards to the magnitude to which avid readers differ from the rest in terms of LD scores, these results support finding of previous studies about the positive correlation between frequent EE reading and language development (Busby 2021: 72; Tam & Reynolds 2022: 19; Warnby 2022: 133).

Next, similar effect sizes and 95% CI can be observed for the writers and self-talkers. These two groups are also alike in terms of size. Cohen's d shows medium effect sizes for the writers (d = .622) and the self-talkers (d = .658). It can be assumed that the true mean does lie in the range of medium effect sizes, as the confidence intervals of both groups covers almost the full range of medium effect sizes, and the upper boundary even reaches large effects. Therefore, it can be argued that students who frequently write or think/talk to themselves in English are able to produce texts with a higher LD compared to other learners.

Table 12: Difference in LD scores between 'niche' group participants and rest of participants

	N	p-value	Cohen's d	95% CI
avid readers	4	p = .021	d = .693	[.262, .894]
the writers	16	p < .001	d = .622	[.385, .782]
the talkers	26	p < .001	d = .542	[.319, .709]
the self-talkers	15	p < .001	d = .658	[.429, .807]

Lastly, participants who frequently engage in speaking activities, i.e. the talkers, also differ considerably from the rest. Even though Cohen's d is smaller than for the other groups (d = .542), it is still well within medium effect sizes. The 95% CI is similar to the writers / self-talkers in terms of range but reaches more into smaller effect sizes. It could be argued that the talkers are not really a niche group, as its 26 members constitute more than a third of participants. However, when looking at the activities the talkers engage in, this can be explained by the fact that frequent engagement is

distributed between the three speaking activities. Rarely does one member engage in more than one speaking activity on a frequent basis, and not one activity is highly favoured over the other. Additionally, outside of members of this group, the rest of participants indicated very low engagement in speaking activities, which explains while those were among the most unpopular activities outlined above.

There is one student in this study who is a member of all four niche groups discussed above and whose texts exhibit by far the highest MATTR scores (LD mean of both texts = .800, LD descriptive text = .774, LD narrative text = .826). She outscored the next participant in terms of mean LD score by .087. This high level of LD can be explained through her frequently engaging in niche activities, such as frequent reading, writing, or talking to other people in English. Yet, her overall EE engagement is also very high (Mdn_{EE} = 4). She listens to music, watches movies and series on a daily basis, and plays video games, reads on social media, and writes on social media on a weekly basis. There are very few EE activities she rarely does or doesn't do at all, and those are examples of endeavours that are rather specialised, such as reading and writing comics/or mangas, or activities almost all students are not involved with, such as watching movies in the cinema or listening to audiobooks. Thus, the high LD of her texts can't necessarily be explained by either her engaging in qualitative EE activities that required higher language processing or an overall high frequency of engagement, but by a mixture of both. While EE engagement does provide a rationale for this participant being able to produce texts with a high LD, her demographic and linguistic background might also have an influence. She was born in Vietnam and exclusively speaks Vietnamese with her family members. Therefore, German plays an insignificant role in her life, which is why she mostly uses English to engage with friends. She is one of the few participants who indicated English as the number one language she uses in her free time. Thus, her background dictates to some extent the linguistic patterns of her leisure time activities.

Revisiting the discussion of quantity versus quality of EE and their relationship to LD, the results point towards a higher influence of niche activities on LD in writing. Effect sizes across all niche groups show a higher magnitude in terms of correlation with LD scores than EE median scores. Avid readers stand out with the highest effect size; however, the wide CI points to some uncertainty about the reliability of this result. Furthermore, activities such as reading books or other English texts show only small positive correlations with LD. Thus, to what extent frequent reading in English benefits LD in writing can not necessarily be ascertained. LD scores of the writers, self-talkers, and talker, also differ from other participants to a medium effect but contrary to the avid readers, CI suggest a higher

reliability of effect sizes. It can therefore be argued that, while quantity of EE engagement matters, it is activities that required a higher level of language processing, such as writing, speaking, and self-directed speech, that have the most substantial impact on lexical diversity in writing.

7 Discussion of results

As outlined in the beginning, this thesis aims at contributing to the small body of research about EE and its influence on writing. It offers insights into EE practices of lower-secondary Viennese students and how those practices correlate with LD scores of their writing performance. In the following chapter, the most crucial findings are discussed and interpreted, with a particular emphasis on how they align with or diverge from results of previous research in this field and how the study's limitations restrict the generalisation of results to other contexts.

First of all, the mapping of students' EE practices revealed several similarities to previous research. Participants exhibit a wide range in terms of how frequently they engage in EE, and they reported a lower quantity of EE than studies with older learners. It has to be noted here that the questionnaire used in this study is not without its limitation. First of all, it does not track the exact time students spend on EE, which makes comparisons to some other studies on EE not feasible. Second, the sample size is rather small; therefore, limiting the potential for generalisation of the results to contexts other than this thesis. Lastly, as EE is something students do in their private free time, this study had to rely on a questionnaire where students report their EE habits themselves. Direct observations, however, are not feasible in research on EE, due to time constraints and reasons of privacy. Thus, this study had to rely on self-reported data. Future research could gather qualitative data, e.g. interviews, as a complement to quantitative data to get a more reliable picture of students' EE engagement.

When it comes to daily engagement, in previous research in a context similar to this study, Schwarz (2020: 212) found that her upper-secondary participants engage on average in 9.43 daily EE activities and 96.50% of participants do at least one daily activity. In comparison, students in this study reported a mean of 8.03 daily activities and 82.43% of learners do at least one EE activity on a daily basis. Furthermore, 13 participants in this study indicated no daily EE engagement, compared to seven learners in Schwarz' study (2020: 212). This gap in frequency of engagement between younger and older learners becomes even more apparent when looking at weekly EE engagement. Schwarz

(2020: 212) reports a mean of 16.87 weekly activities, compared to students in this study engaging on average in 6.57 weekly activities. The majority of learners (52.70%) in this study also reported very low EE median scores of 1 or 2. Thus, it can be assumed that EE plays a more important role in the daily lives of older learners and that English has littles relevance for some younger learners' free time activities in this study.

Regarding specific EE activities students engage in most frequently, the results in this study are in line with previous research. Before conducting the data analysis, it was presumed that the majority of students in this study engage in EE activities out of a want for entertainment without an explicit intent to learn English. The result support this hypothesis, as activities focused on entertainment are among the most popular EE activities. Playing video games is the most favoured activity of participants, followed by listening to music, watching movies/series/videos, and reading on social media. Other studies reported the same most frequent activities, albeit the ranking might sometimes differ (e.g. de Wilde, Brysbeart & Eyckmans 2020; Peters 2018, Schwarz 2020, Tam & Reynolds 2022, Warnby 2022). These results also reflect those of previous research on age differences, namely that playing video games seems to be an EE activity that is especially popular with younger learners (Puimège & Peters 2019: 965). In her study on older learners, Schwarz (2020) found that her participants play video games as a general leisure time as well as EE activity to a less frequent extent. How students spend their free time seems to shift with age and EE activities might change accordingly. This assumption is also supported by the fact that, in this study and previous research, the most popular EE activities align to a certain degree with the general leisure time activities participants frequently engage in (Schwarz 2020: 213). Overall, activities carried out online have a high relevance in how participants spend their free time, as the most popular activities have some connection to the internet or digital media. Consequently, analogue media such as books play a diminished role in students' daily lives. Gender-specific differences with regards to most popular EE activities found in previous studies are also confirmed by the results of this thesis (Hannibal Jensen 2017; Olsson & Slyvén 2015). Male students prefer to play video games while female learners listen to music the most. When it comes to quantity of EE, contrary to previous research, boys do not engage more frequently in EE than girls. The analysis showed that female students reported a higher frequency of engagement. However, the correlation analysis indicated only a small effect sizes for the difference between gender and the p-value was just above the threshold of statistical significance. What can be ascertained form these results, however, is that male learners in this study do not engage more frequently than female students.

This study supports evidence from previous observations about least popular EE activities. Overall, as in studies by Schwarz (2020: 213), Sundqvist (2009: 117), and Warnby (2022), participants reported to be less engaged in productive EE activities and more in receptive activities that serve an entertainment purpose. Especially writing and reading are rarely done by students in this study and in previous research (Avello et al. 2019: 84; Cabot 2016: 173; de Wilde, Brysbeart & Eyckmans 2020: 177; Inaba 2018: 92; Olsson & Sylvén 2015: 87). Both activities seem to be limited in an EE context to shorter texts in connection with social media. Olsson and Sylvén (2015: 87) found similar results for EE writing and Schwarz (2020: 213) for reading. Extensive reading and especially reading books seems to be an activity students shy away from in EE, which might be connected to the fact that most popular free time activities have some connection to the internet. Activities such as extensive reading or writing may also reflect students' motivations when engaging with English in an EE context. They often require more effort and may be perceived as too similar to academic work; thus, they offer less appeal in learners free time when they want to relax.

Second, this thesis set out to find possible correlations between participants' EE practices and the LD of their texts. Before exploring the influence of EE on LD in writing, MATTR scores, as a measure of LD, were calculated. Due to this study being a MEd thesis with constraints on access to participants and time resources, only two text per participant were collected. Having a greater data sample with more extensive language samples from participants would yield more valid results. Furthermore, although this study employs a measure of LD that is not influenced by text length, more extensive language samples would provide a more reliable measure of LD and more opportunities for participants to show their productive vocabulary. Another limitation with regards to how LD was measured in this study is that the MATTR only aims at capturing one property of LD, namely variability. Even though variability was shown to be a reliable indicator of LD and writing proficiency (Crossley et al. 2014: 10; Vögelin et al. 2019: 59), using multiple measures that aim at different properties of LD would provide a more complete picture of students LD in writing. This would certainly increase the complexity of the data analysis and interpretation and would go beyond the scope of a master thesis.

The results show that students exhibit a wide range of scores, with some reaching high MATTR scores normally observed with older, higher-level learners and others only demonstrating basic level of LD. Yet, overall, participants in this study exhibit a lower mean MATTR score than students in previous studies by Kyle, Crossley, and Jarvis (2021) and Zenker and Kyle (2021). However, learners in these studies were older university students with a higher language level. The narrative text displayed the

most diverse vocabulary and students produced lower LD scores for the descriptive text. This suggests that the text genre might have an influence on how lexically diverse students are able to write. When looking at group differences, the analysis revealed that 'AHS' students were able to produce texts with a higher LD than 'Standard' students. Even though the distinction between 'AHS' and 'Standard' in Austrian 'Mittelschulen' cannot readily used as an indicator of overall language proficiency, it does reflect to a certain extent how students performed in class and assessments and points to a potentially higher language level. Thus, this difference between the two groups might be an indicator of students with a higher language level being able to write more lexically diverse texts. Lastly, LD scores of boys and girls were also investigated with regards to potential differences. The analysis, however, showed no statistically significant impact of gender on LD scores.

The review of studies showed that the quantity of EE positively correlates with certain language skills, but the quality of EE practices, with students engaging in activities that require language processing on a higher level, trumps sheer frequency of engagement. The analysis in this study showed that quantity of EE, expressed through an EE median score, positively correlates with LD scores to a small to medium effect (τ = .333). Furthermore, when grouping students according to frequency of engagement, the high EE group differs from the low EE group to medium effect. When comparing these findings to correlation analysis done for specific EE activities, similar effect sizes can be observed for some activities but most only exhibit negligible effect sizes or the correlation does not reach statistical significance. 'Speaking face to face' positively correlates with LD scores to the highest effect comparable to the small to medium effect of EE median scores (τ = .379), followed by 'watching video on the internet' (τ = .345), and the summary variable 'reading on social media' (τ = .330). The analysis of other variables that also showed statistically significant positive correlation only revealed small effect sizes lower than the EE median score. While several studies showed that gaming has a positive effect on receptive vocabulary knowledge (Busby 2021: 72; de Wilde, Brysbeart & Eyckmans 2020: 178; Hannibal Jensen 2017: 13; Puimège & Peters 2019: 966; Sundqvist 2009: 156; Sundqvist 2019: 95; Warnby 2022: 133), it does not have a statistically significant correlation with LD scores in this study. Thus, although gaming potentially improves learners' receptive vocabulary, it does not influence how well learners are able to use vocabulary productively in writing in the context of this study. Overall, these results do not provide a clear indication whether quantity or specific EE activities have a more pronounced effect on LD. However, when looking at niche activities and students who frequently engage in those activities, it can be observed that these learners differ from the rest of participants to medium to large effect sizes. Especially those who

engage in activities that require active language use, such as writing, speaking, or self-directed speech, produced texts with considerably higher LD scores than the rest of students. Therefore, consistent with the literature, these results support the assumption that, while frequency of engagement has an influence on the development of language skills, engaging in high quality EE activities correlates with language skills to a higher effect than quantity of engagement (de Wilde, Brysbeart & Eyckmans 2020: 180).

To conclude, results in this thesis mostly correspond with previous findings in research on EE. The mapping of EE practices revealed that the younger participants in this study differ from older learners in terms of quantity as well as quality of EE. Frequency of engagement as well as specific activities showed a positive correlation with MATTR scores; yet, the most pronounced effects could be observed with 'niche' activities, supporting arguments by previous studies about higher quality activities correlating with language skills to the highest extent.

8 Conclusion

This study set out to find potential correlations between EE practices of Viennese lower-secondary students and their LD in writing. While the majority of research on EE investigated the relationship between EE and receptive vocabulary, this study chose to focus on writing in order to fill the lack of research regarding EE and this particular language skill. Two main aims were followed in this study. First, EE practices of students were mapped to answer questions about the quality and quantity of their EE engagement. Second, the main research question focused on whether EE correlates with LD scores in students' writing.

The findings revealed that EE activities for entertainment, such as playing video games, listening to music, and watching audiovisual input, were the most popular among learners, with receptive activities being favoured over productive ones. In contract, writing and reading, especially extensive reading, were reported as the least favoured activities. When it comes to total frequency of engagement, this study's younger participants engaged in EE to a lesser extent than older students of previous studies (Schwarz 2020). There were also some gender differences, with female participants engaging more frequently in EE overall and showing a preference for listening to music, while male learners favoured playing video games. In terms of quantity, in contract to previous research, this study did not find any statistically significant differences between boys and girls.

In addressing the main research question, the results showed that frequency of EE as well as specific activities positively correlate with LD scores of students' texts, meaning that the more students are engage with English or the more students engage in activities of a higher quality in terms of language processing, the higher the LD their texts exhibit. While most EE activities showed lower positive correlations with LD scores than the EE median scores, an indicator of quantity of EE, niche activities only few students engage in, such as extensive reading, writing, speaking, and self-talk, exhibited the highest positive correlation with LD scores. Even though quantity of EE was shown to have positive benefits for LD in writing, this emphasizes the importance of engaging in productive and high-quality EE activities to enhance writing proficiency.

These results have the potential to improve language teaching in the classroom and help students to be more engaged in English lessons. First, research in this field helps us to gain an understanding of how learners engage with English outside the classroom. This allows teachers to more closely align their existing classroom practices with language and activities students encounter in an EE context. Connecting language learning inside the classroom with authentic input students receive in their free time has the potential to increase motivation and engagement in class. Teachers should be aware of the benefits of EE, so that they can advise students on how to improve their English skills outside the classroom. It has to be noted that, given limited resources and time, catering to students' individual EE practices in language teaching might not be feasible. However, this study and others have shown that some EE activities are popular across age groups, genders, and individual learners, and these activities, e.g. 'listening to music', could easily be utilised in different school-related tasks such as homework or different projects.

Second, by having a sound understanding of EE and recognising its role as a relevant complement to traditional classroom learning, teachers will be able to better promote informal language learning outside school. This approach could encourage students to view language learning as a continuous process that does not stop at the door of the classroom, integrating EE practices into their daily lives. By shifting this perspective, students potentially view language learning as something dynamic and relevant to their personal interest and free time pursuits. Several studies suggest that time spent on EE far outreaches time in instructed English lesson (e.g Olsson & Sylvén 2015; Schwarz 2020). This time could then be used more towards improving students' language skills. If students use this time not only for entertainment purposes but also employ it so they can informally develop their language skills, the benefits for their overall language proficiency should be considerable.

Future research in the field of EE and writing could help expand on these results and curb some of the limitations of this study mentioned in the preceding chapter. As already mentioned, compiling a larger data sample would allow for more generalisations. In terms of mapping learners EE practices, a mixed method approach might bring about more reliable results. Moreover, investigating correlations of EE with different text genre and comparing these correlations would yield interesting results about which genres are influenced by EE and how this might influence teaching writing at school. In general, more research should be done on how EE influences language skills aside from receptive vocabulary teaching. Given that 'speaking face to face' had a pronounced effect on writing in this thesis, it would be interesting to examine how EE correlates with overall speaking proficiency or different speaking skills.

9 References

Avella, Maria; Camacho-Miñano, Maria-del-Mar; Urquia-Grand, Elena; del Campo, Cristina. 2019. "Do you use English in your daily life?: Undergraduate students' perceptions of their extramural use of English". *Journal of Teaching in International Business* 30(1), 77-94.

Bauer, Laurie; Nation, Ian Stephen Paul. 1993. "Word families". *International Journal of Lexicography* 6(4), 253-279.

Benson, Phil. 2011. "Language learning and teaching beyond the classroom: An introduction to the field". In Benson, Phil; Reinders, Hayo (eds.). 2011. *Beyond the language classroom*. Basingstoke: Palgrave Macmillan, 7-16.

Bereiter, Carl; Scardamalia, Marlene. 1987. *The psychology of written composition*. Hillsdale, NJ: Erlbaum.

BMBWF. 2024a. *Lehrplan der Volksschule*. https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10009275 (28 Jul. 2024).

BMBWF. 2024b. *Lehrplan der Mittelschule*. https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20007850 (28 Jul. 2024).

Bollansée, Lieven; Puimège, Eva; Peters, Elke. 2020. "Watch out! Behind you is the enemy!". In Werner, Valentin; Tegge, Friederike (eds.). *Pop culture in language education: Theory, research, practice.* London: Rouledge, 199-214.

Brown, Dale; Stoeckl, Tim; McLean, Stuart; Stewart, Jeff. 2022. "The most appropriate lexical unit for L2 vocabulary research and pedagogy: A brief review of the evidence". *Applied Linguistics* 43, 596-602.

Bulté, Bram; Housen, Alex. 2014. "Conceptualizing and measuring short-term changes in L2 writing complexity". *Journal of Second Language Writing* 26(2014), 42-65.

Busby, Nicole Louise. 2021. "Words from where?: Predictors of L2 English vocabulary among Norwegian university students". *International Journal of Applied Linguistics* 172(1), 58-84.

Cabot, Michel. 2016. "In or out of school?: Meaningful output with digital and non-digital artefacts within personal English". *Nordic Journal of Digital Literacy* 11, 165-184.

Cohen, Jacob. 2013. *Statistical power analysis for the behavioural sciences.* Burlington: Elsevier Science.

Cole, Jason; Vanderplank, Robert. 2016. "Comparing autonomous and class-based learners in Brazil: Evidence for the present-day advantages of informal, out-of-class learning". *System* 61, 31-42.

Cohen, Louis; Manion, Lawrence; Morrison, Keith. 2018. *Research methods in education.* London: Routledge.

Council of Europe. 2020. *Common European framework of reference for languages*. https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4 (28 Jul. 2024).

Covington, Michael; McFall, Joe. 2010. “Cutting the Gordian knot: The moving-average type-token ratio (MATTR)”. *Journal of Quantitative Linguistics* 17, 94-100.

Crossley, Scott; Salsbury, Tom; McNamara, Danielle. 2012. “Predicting the proficiency level of language learners using lexical indices”. *Language Testing* 29(2), 243-263.

Crossley, Scott; Kyle, Kristopher; Allen, Laura; Guo, Liang; McNamara, Danielle. 2014. „Linguistic microfeatures to predict L2 writing proficiency”. *Journal of Writing Assessment* 7, 1-23.

De Wilde, Vanessa; Eyckmans, June. 2017. “Game on!: Young learners’ incidental language learning of English prior to instruction”. *Studies in second language learning and teaching* 7, 673-694.

De Wilde, Vanessa; Brysbeart, Marc; Eyckmans, June. 2020. “Learning English through out-of-school exposure: Which levels of language proficiency are attained and which types of input are important?”. *Bilingualism* 23(1), 171-185.

Dohmen, Günther. 2001. *Das informelle Lernen: Die internationale Erschließung einer bisher vernachlässigten Grundform menschlichen Lernens für das lebenslange Lernen aller.* Bonn: Bundesministerium für Bildung und Forschung.

Ellis, Rod. 2009. “Implicit and explicit learning, knowledge and instruction”. In Ellis, Rod; Loewen, Shawn; Elder, Catherine; Reinders, Hayo; Erlam, Rosemary; Philip, Jenefer (eds.). *Implicit and explicit knowledge in second language learning, testing and teaching.* Bristol: Multilingual Matters, 3-25.

Field, Andy. 2012. *Discovering statistics using R.* London: SAGE Publications.

Flower, Linda; Hayes, John. 1981. “A cognitive process theory of writing”. *College Composition and Communication* 32, 365-387.

Francis, Winthrop; Kučera, Henry. 1982. *Frequency analysis of English usage: Lexicon and grammar.* Boston, MA: Houghton Mifflin.

Gardner, Dee. 2007. “Validating the construct of word in applied corpus-based vocabulary research: A critical survey”. *Applied Linguistics* 28, 241-265.

González, Melanie. 2017. “The contribution of lexical diversity to college-level writing”. *TESOL Journal 8(4), 899-919.*

Gould, Stephen Jay. 1989. *Wonderful life: The Burgess Shale and the nature of history*. New York, NY: Norton.

Grabe, William; Kaplan, Robert. 1996. *Theory and practice of writing: An applied linguistic perspective*. London: Longman.

Hahn, Magdalena. 2017. “Extramural vocabulary acquisition: A survey of students in Viennese business middle schools (HAS)”. MEd Thesis, Department of English and American Studies, University of Vienna.

Hahn, Magdalena. 2018. “Extramural vocabulary acquisition: A survey of students in Viennese business middle schools (HAS)”. *VIEWS* 27 (2018), 1-21.

Hannibal Jensen, Signe. 2017. “Gaming as an English language learning resource among young children in Denmark”. *Calico Journal* 34(1), 1-19.

Heaps, Harold Stanley. 1978. *Information retrieval: Computatuinal and theoretical aspects.* New York, NY: Academic Press.

Henry, Alastair; Korp, Helena; Sundqvist, Pia; Thorsen, Cecilia. 2018. "Motivational strategies and the reframing of English: Activity design and challenges for teachers". *TESOL Quarterly* 52(2), 247-273.

Housen, Alex; Kuiken, Folkert; Vedder, Ineke. 2012. "Complexity, accuracy, fluency". In House, Alex (ed.). *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA.* Amsterdam: John Benjamins Publishing Company, 1-20.

Hyland, Ken. 2022. *Teaching and researching writing*. New York, NY: Routledge.

Inaba, Miho. 2018. *Second language literacy practices and language learning outside the classroom.* Bristol: Multilingual Matters.

Jarvis, Scott. 2013a. "Defining lexical diversity". In Jarvis, Scott; Daller, Michael (eds.). *Vocabulary knowledge: Human ratings and automated measures.* Amsterdam: John Benjamins Publishing Company, 13-43.

Jarvis, Scott. 2013b. "Capturing the diversity in lexical diversity". *Language Learning* 63, 87-106.

Jarvis, Scott. 2017. "Grounding lexical diversity in human judgments". *Language Testing* 34(4), 537-553.

JASP Team. 2024. *JASP (Version 0.19.0).* [Computer Program].

Jurkovič, Violeta. 2019. "Online informal learning of English through smartphones in Slovenia". *System* 80, 27-37.

Kojima, Masumi; In'nami, Yo; Kaneta, Taku. 2022. "L2 writing and its external correlates: A meta-analysis". In Jeon, Eun Hee; In'nami, Yo (eds.). *Understanding L2 proficiency: Theoretical and meta-analytic investigations.* Amsterdam: John Benjamins Publishing Company, 159-212.

Kojima, Masumi; Kaneta, Taku. 2022. "L2 writing and its internal correlates". In Jeon, Eun Hee; In'nami, Yo (eds.). *Understanding L2 proficiency: Theoretical and meta-analytic investigations.* Amsterdam: John Benjamins Publishing Company, 109-158.

Kyle, Kristopher; Crossley, Scott; Jarvis, Scott (2021). "Assessing the validity of lexical diversity using direct judgements". *Language Assessment Quarterly 18*(2), 154-170.

Lai, Chun; Zhu, Weimin; Gong, Gang. 2015. "Understanding the quality of out-of-class English learning". *TESOL Quarterly* 49(2), 278-308.

Laufer, Batia. 2005. "Focus on form in second language vocabulary learning". *Eurosla yearbook* 5, 223-250.

Laufer, Batia; Hulstijn, Jan. 2001. "Incidental vocabulary acquisition in a second language: The construct of task-induced involvement". *Applied Linguistic* 22, 1-26.

Lee, Ju Seong. 2019a. "Informal digital learning of English and second language vocabulary outcomes: Can quantity conquer quality?". *British Journal of Educational Technology* 50(2), 767-778.

Lee, Ju Seong; Dressman, Mark. 2018. "When IDLE hands make an English workshop: Informal digital learning of English and language". *TESOL Quarterly* 53(2), 435-445.

Lee, Ju Seong; Sylvén, Liss Kerstin. 2021. "The role of informal digital learning of English in Korean and Swedish EFL learners' communication behaviour". *British Journal of Educational Technology* 52, 1279-1296.

Leona, Nihayra; van Koert, Margreet; van der Molen, Maurits; Rispens, Judith; Tijms, Jurgen; Snellings, Patrick. 2021. "Explaining individual differences in young English language learners' vocabulary knowledge: The role of extramural English exposure and motivation". *System* 96, 102402.

Liu, Zhibo; Dou, Juhua. 2023. "Lexical density, lexical diversity, and lexical sophistication in simultaneously interpreted texts: A cognitive perspective". *Frontiers in Psychology* 14, 1-11.

Lo Bianco, Joseph. 2014. "Domesticating the foreign: Globalization's effects on place/s of languages". *The Modern Language Journal* 98(1), 312-325.

Maamuujav, Undarmaa; Olson, Carol Booth; Chung, Huy. 2021. "Syntactic and lexical features of adolescent L2 students' academic writing". *Journal of Second Language Writing* 53(2021), 1-16.

Mair, Christian. 2020. "English in the German-speaking world: An inevitable presence". In Hickey, Raymond (ed.). *English in the German-speaking world.* Cambridge: Cambridge University Press, 13-30.

McCarthy, Philip. 2005. "An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity (MTLD)". PhD thesis, University of Memphis.

McCarthy, Philip; Jarvis, Scott. 2007. "Vocd: A theoretical and empirical evaluation". *Language Testing* 24(4), 459-488.

McLean, Stuart. 2018. "Evidence for the adoption of the flemma as an appropriate word counting unit". *Applied Linguistics* 39, 823-845.

Meara, Paul. 1990. "A note one passive vocabulary". *Second Language Research* 6, 150-154.

Melka, Francine. 1997. "Receptive vs. productive aspects of vocabulary". In Schmitt, Norbert; McCarthy, Michael (eds.). *Vocabulary: Description, acquisition and pedagogy.* Cambridge: Cambridge University Press, 84-102.

Milton, James; Fitzpatrick, Tee. 2014. *Dimensions of vocabulary knowledge.* Basingstoke: Palgrave Macmillan.

Nation, Ian Stephen Paul. 2001. *Learning vocabulary in another language.* (1st edition). Cambridge: Cambridge University Press.

Nation, Ian Stephen Paul. 2022. *Learning vocabulary in another language.* (3rd edition). Cambridge: Cambridge University Press.

Olsson, Eva. 2016. "On the impact of extramural English and CLIL on productive vocabulary". PhD thesis, Department of Education and Special Education, University of Gothenburg.

Olsson, Eva; Sylvén, Liss Kerstin. 2015. "Extramural English and academic vocabulary: A longitudinal study of CLIL and non-CLIL students in Sweden". *Journal of Applied Language Studies* 9(2), 77-103.

Peters, Elke. 2018. "The effect of out-of-class exposure to English language media on learners' vocabulary knowledge". *International Journal of Applied Linguistics* 169(1), 142-168.

Peters, Elke; Webb, Stuart. 2018. “Incidental vocabulary acquisition through viewing L2 television and factors that affect learning”. *Studies in Second Language Acquisition* 40(3), 551-577.

Peters, Elke; Noreillie, Ann-Sophie; Heylen, Kris; Bulté, Bram; Desmet, Piet. 2019. “The impact of instruction and out-of-school exposure to foreign language input on learners’ vocabulary knowledge in two languages”. *Language Learning* 69(3), 747-782.

Plonsky, Luke; Oswald, Frederick. 2014. “How big is big? Interpreting effect sizes in L2 research”. *Language Learning* 64(4), 878-912.

Puimège, Eva; Peters, Elke. 2019. “Learners‘ English vocabulary knowledge prior to formal instruction: The role of learner-related and word-related variable”. *Language Learning* 69(4), 943-977.

Reinders, Hayo; Benson, Phil. 2017. “Research agenda: Language learning beyond the classroom”. *Language Teaching* 50, 561-578.

Roca de Larios, Julio; Manchón, Rose; Murphy, Liz; Marín, Javier. 2008. “The foreign language writer’s strategic behaviour in the allocation of time to writing processes”. *Journal of Second Language Writing* 17, 30-47.

Sasaki, Miyuki. 2000. “Toward an empirical model of EFL writing processes: An exploratory study”. *Journal of Second Language Writing* 9, 259-291.

Schmidt, Richard. 1994. ”Deconstructing consciousness in search of useful definitions for applied linguistics”. *AILA Review* 11, 11-26.

Schmitt, Nobert. 2010. *Researching vocabulary: A vocabulary research manual.* Basingstoke: Palgrave Macmillan.

Schoonen, Rob. 2022. “L2 writing: Theory and research”. In Jeon, Eun; In’nami, Yo (eds.). *Understanding L2 proficiency: Theoretical and meta-analytic investigations*. Amsterdam: John Benjamin Publishing Company, 87-108.

Schurz, Alexandra; Coumel, Marion; Hüttner, Julia. „Accuracy and fluency teaching and the role of extramural English: A tale of three countries”. *Languages* 7(1), 35.

Schwarz, Marlene. 2020. “Beyond the walls: A mixed methods study of teenagers’ extramural English practices and their vocabulary knowledge”. PhD thesis, Department of English and American Studies, University of Vienna.

Segalowitz, Norman. 2016. “Second language fluency and its underlying cognitive and social determinants”. *International Review of Applied Linguistics in Language Teaching* 54(2), 79-95.

Silva, Renita; Clahsen, Harald. 2008. “Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English”. *Bilingualism* 11, 245-260.

Sockett, Geoffrey. 2014. *The online informal learning of English.* (1st edition). Basingstoke: Palgrave Macmillan.

Stubbs, Michael. 2001. *Words and phrases: Corpus studies of lexical semantics.* Oxford: Blackwell Publishing.

Sun, Ting; Nam, Hyunjeong. 2023. “Variations in vocabulary depth among second language learners with comparable vocabulary breadths: An empirical investigation”. *Journal of Psycholinguistic Research* 52, 2473-2498.

Sundqvist, Pia. 2009. “Extramural English matters: Out-of-school English and its impact on Swedish ninth graders’ oral proficiency and vocabulary”. PhD thesis, Faculty of Arts and Education English, Karlstad University.

Sundqvist, Pia. 2019. “Commercial-off-the-shelf games in the digital wild and L2 learner vocabulary”. *Language Learning & Technology* 23(1), 87-113.

Sundqvist, Pia; Sylvén, Liss Kerstin. 2016. *Extramural English in teaching and learning: From theory and research to practice.* London: Palgrave Macmillan.

Tam, Ho; Reynolds, Barry Lee. 2022. “The relationship between extramural English engagement and the vocabulary size of L1 Cantonese speakers in Macau”. *International Journal of Applied Linguistics,* DOI: 10.1075/itl.21003.tam.

Treffers-Daller, Jeanine; Parslow, Patrick; Williams, Shirley. 2018. “Back to basics: How measures of lexical diversity can help discriminate between CEFR levels”. *Applied Linguistics* 39(3), 302-327.

Verspoor, Marjolijn; de Bot, Kees; Eva, van Rein. 2011. “English as a foreign language”.In de Houwer, Annick; Wilton, Antje (eds.). *English in Europe today.* Philadelphia, PA: John Benjamins Publishing Company, 147-169.

Verspoor, Marjolijn; Schmid, Monika; Xu, Xiaoyan. 2012. “A dynamic usage based perspective on L2 writing”. *Journal of Second Language Writing* 21(3), 239-263.

Vögelin, Cristina; Jansen, Thorben; Keller, Stefan; Machts, Nils; Möller, Jens. 2019. ”The influence of lexical features on teacher judgements of ESL argumentative essays”. *Assessing Writing* 39, 50-63.

Wang, Xuan. 2014. “The relationship between lexical diversity and EFL writing proficiency”. *University of Sydney Papers in TESOL* 9(2014), 65-88.

Warnby, Marcus. 2022. “Receptive academic vocabulary knowledge and extramural English involvement: Is there a correlation?”. *International Journal of Applied Linguistics* 173(1), 120-152.

Webb, Stuart. 2010. “A corpus driven study of the potential for vocabulary learning through watching movies”. *International Journal of Corpus Linguistics* 15, 497-519.

Wingate, Ursula; Harper, Rowena. 2021. “A case study of the writing processes of a successful and an unsuccessful student”. *Journal of English for Academic Purposes* 49, 100848.

Yu, Guoxing. 2010. “Lexical diversity in writing and speaking task performance”. *Applied Linguistics 31(2), 236-259.*

Yule, Udny. 1944. *The statistical study of literary vocabulary*. Cambridge: Cambridge University Press.

Zenker, Fred; Kyle, Kristopher. 2021. “Investigating minimum text lengths for lexical diversity indices“. *Assessing Writing* 47, 100505.

10 Appendix

Englisch in deiner Freizeit

Liebe Schülerin, lieber Schüler,

Um mein Studium an der Universität Wien zu beenden, schreibe ich eine große Arbeit über Englisch in der Freizeit. Diese Arbeit wird mir hoffentlich helfen, den Englischunterricht besser an dich anzupassen. Du brauchst ca. 20 min zum Ausfüllen. Deine Antworten sind **anonym** und werden streng vertraulich behandelt. Dieser Fragebogen ist kein Test, darum gibt es auch keine richtigen oder falschen Antworten. Bitte beantworte die Fragen **alleine, spontan** und **ehrlich**.

Danke, dass du dir Zeit nimmst mir bei meiner Arbeit zu helfen!

Du bekommst einen persönlichen Code, damit ich deinen Fragebogen anderen Teilen der Studie zuordnen kann,

Mein Code: ☐☐ 2. und 3. Buchstabe deines Vornamens
☐☐ Geburtsmonat (z.B. 05 für Mai)

So füllst du den Fragebogen richtig aus:

Meistens musst du dich für eine Antwort entscheiden und …

… ein Kästchen ankreuzen — Hast du Geschwister? ☒ Ja ☐ Nein

… die zutreffende Zahl ankreuzen

Wie oft isst du Erdbeeren?	Fast nie	Selten	Oft	Sehr oft
	1	~~2~~	3	4

Bei allen Fragen, die einen ➡ haben, kannst du deine Antwort in eigenen Worten geben

Bitte schreibe leserlich! — Wie findest du Clowns? ➡ *Ich finde Clowns witzig.*

Fall du Fragen hast, frag mich einfach 😊

1. Informationen über dich

Hier möchte ich ein paar allgemeine Dinge über dich wissen, die mir bei beim Auswerten deines Fragebogens helfen könnten.

➢ *Dein Fragebogen bleibt trotzdem anonym und deine Antworten verraten mir nicht, wer du bist!*

1a. dein Geburtsjahr: __ __ __ __

1b. In welchem Land bist du geboren? _____

1c. dein Geschlecht: ☐ weiblich ☐ männlich ☐ divers ☐ möchte ich nicht sagen

1d. In welcher Gruppe bist du in Englisch? ☐ Standard ☐ AHS

1d. Welche Sprache sprichst du … *Wenn eine Frage nicht auf dich zutrifft, lass die Zeile daneben einfach leer!*	⬇ *Bitte trag hier die Sprachen ein!*
…mit deiner Mutter?	
…mit deinem Vater?	
…mit deinen Geschwistern?	
…mit anderen Menschen bei dir zu Hause (z.B. Großeltern, Partner oder Partnerin der Mutter oder Vater)	
…meistens mit deinen besten Freunden?	
Welche Sprache sprechen deine Eltern miteinander?	

1e. Welche Sprachen verwendest du in deiner FREIZEIT (Freunde, Familie, Aktivitäten,…)**?**

1. ➡
2. ➡
3. ➡
4. ➡

1f. Gibt es Menschen, mit denen du (außerhalb des Schulunterrichts) regelmäßig Englisch sprichst? ☐ Ja ☐ Nein

Wenn **ja**, mit wem? ➡ _____

2. Deine technische Ausstattung

Mich interessiert, welche Geräte du in deiner Freizeit zu Verfügung hast.

2a. Gibt es bei dir zu Hause einen Internetzugang? ☐ Ja ☐ Nein ☐ nur manchmal

2b. Welche der folgenden Dinge habt ihr zu Hause?

☐ Handy	☐ Fernseher
☐ Computer (Laptop, Standcomputer, Netbook)	☐ E-Book-Reader (z.B. Kindle)
☐ Tablet / iPad	☐ Radio / Musikanlage
☐ Spielekonsole (z.B. Playstation, Wii, Nintendo)	andere Geräte: ➡ _____

3. Was machst du in deiner Freizeit?

Hier geht es allgemein um Aktivitäten, die du in deiner Freizeit machst. In welcher Sprache du diese Sachen machst ist hier noch nicht wichtig!

3a. Wie oft machst du diese Tätigkeiten in deiner Freizeit <u>allgemein</u>? ➢ *Bitte fülle auch Fragen aus, wo du eine Antwort selbst schreiben musst!*	**(Fast) nie**	**Ein paar Mal pro Jahr**	**Ein paar Mal pro Monat**	**Ein paar Mal pro Woche**	**(Fast) täglich**
Musik hören	1	2	3	4	5
Fernsehen	1	2	3	4	5
Filme oder Serien ansehen	1	2	3	4	5
Videos im Internet ansehen	1	2	3	4	5
Soziale Netzwerke benutzen (z.B. TikTok, Instagram, Facebook,…)	1	2	3	4	5
Spiele alleine spielen (Spielekonsolen, Laptop, Computer,…)	1	2	3	4	5
Spiele mit anderen gemeinsam spielen (z.b. Multiplayer Online Games)	1	2	3	4	5
Bücher oder E-Books lesen	1	2	3	4	5
Zeitungen oder Zeitschriften lesen (auch online)	1	2	3	4	5
Hörbücher hören	1	2	3	4	5
Freunde oder Freundinnen treffen	1	2	3	4	5
Musik machen (singen oder ein Instrument spielen)	1	2	3	4	5
ins Kino gehen	1	2	3	4	5
andere Aktivitäten *(kreuze auch an wie oft du diese machst!)*: ➡ _____	1	2	3	4	5

4. Englische Freizeitaktivitäten

Mich interessiert was du in deiner Freizeit (=außerhalb des Schulunterrichts) auf Englisch machst

- *Es geht hier nur um Sachen mit Englisch in deiner Freizeit, die du nicht für den Schulunterricht machen musst (also keine Hausübungen, Referate, Leseaufgaben…)!*

4a. Wie oft machst du diese Tätigkeiten in deiner Freizeit auf <u>Englisch</u>? ➢ *Bitte fülle auch Fragen aus, wo du eine Antwort selbst schreiben musst!*	**(Fast) nie**	**Ein paar Mal pro Jahr**	**Ein paar Mal pro Monat**	**Ein paar Mal pro Woche**	**(Fast) täglich**
englischsprachige Musik					
… am Handy hören	1	2	3	4	5
… auf Spotify oder anderen Musikapps hören	1	2	3	4	5
… Musikvideos im Internet ansehen (z.B. auf YouTube)	1	2	3	4	5
… mitsingen	1	2	3	4	5
… selbst singen (auch Karaoke)	1	2	3	4	5
englischsprachige Songtexte					
… lesen	1	2	3	4	5
… übersetzen oder Übersetzungen lesen	1	2	3	4	5
… schreiben	1	2	3	4	5
englischsprachige Filme					
… im Fernsehen ansehen	1	2	3	4	5
… im Internet ansehen (z.B. YouTube, Netflix, …) **mit** Untertiteln	1	2	3	4	5
… im Internet ansehen (z.B. YouTube, Netflix, …) **ohne** Untertitel	1	2	3	4	5
… im Kino ansehen	1	2	3	4	5
englischsprachige Serien					
… im Fernsehen ansehen	1	2	3	4	5
… im Internet ansehen (z.B. YouTube, Netflix, …) **mit** Untertiteln	1	2	3	4	5
… im Internet ansehen (z.B. YouTube, Netflix, …) **ohne** Untertitel	1	2	3	4	5

Welche Sprache verwendest du bei Untertiteln von Filmen und Serien meistens? *(Falls du Untertitel verwendest)*

➡ _____

englischsprachige Sendungen (z.B. Dokus, Reportagen etc.)					
… im Fernsehen ansehen	1	2	3	4	5
… im Internet ansehen	1	2	3	4	5
englischsprachige Videoclips (z.B. YouTube, TikTok,…)					
… im Internet ansehen	1	2	3	4	5
… selbst machen	1	2	3	4	5
englischsprachige Spiele					
… am Tablet oder Handy spielen	1	2	3	4	5
… am Computer, Laptop, oder auf Spielekonsolen spielen	1	2	3	4	5
… mit anderen über das Internet spielen (Multiplayer Online Games)	1	2	3	4	5
Wenn ich mit anderen im Internet spiele, *(Lass diese Frage aus, wenn du nicht im Internet mit anderen spielst!)*					
… chatte ich auf Englisch	1	2	3	4	5

4a. Wie oft machst du diese Tätigkeiten in deiner Freizeit auf <u>Englisch</u>? ➢ *Bitte fülle auch Fragen aus, wo du eine Antwort selbst schreiben musst!*	**(Fast) nie**	**Ein paar Mal pro Jahr**	**Ein paar Mal pro Monat**	**Ein paar Mal pro Woche**	**(Fast) täglich**
englischsprachige Bücher					
… in Papierform lessen	1	2	3	4	5
… digital (z.B. E-Reader wie Kindle) lesen	1	2	3	4	5
… als Hörbuch anhören	1	2	3	4	5
englischsprachige Texte (z.B. Zeitung, Nachrichten, Anleitungen, kurze Geschichten, etc.)					
… in Papierform lesen	1	2	3	4	5
… im Internet lesen	1	2	3	4	5
… schreiben	1	2	3	4	5
englischsprachige Comics / Mangas					
… lesen	1	2	3	4	5
… schreiben	1	2	3	4	5
in sozialen Netzwerken (z.B Instagram, TikTok, Twitter)					
… englischsprachige Nachrichten lesen	1	2	3	4	5
… englischsprachige Nachrichten schreiben	1	2	3	4	5
… Kommentare / Statusmeldungen auf Englisch lesen	1	2	3	4	5
… Kommentare / Statusmeldungen auf Englisch schreiben	1	2	3	4	5
englischsprachige WhatsApp Nachrichten / SMS					
… lesen	1	2	3	4	5
… schreiben	1	2	3	4	5
englischsprachige E-Mails					
… lesen	1	2	3	4	5
… schreiben	1	2	3	4	5
auf Englisch chatten					
Mit wem schreibst du englisch Nachrichten oder chattest du? *(Bitte keine Namen: Freund, Freundin, Verwandte…)* ➡ _____					
Listen oder Notizen auf Englisch schreiben	1	2	3	4	5
Tagebuch auf Englisch schreiben	1	2	3	4	5
andere Texte, die du schreibst *(kreuze auch an wie oft du diese machst!)*: ➡ _____	1	2	3	4	5
Englisch sprechen					
… über Videochat (z.B. Zoom, Teams, Skype, etc.)	1	2	3	4	5
… am Telefon	1	2	3	4	5
… persönlich	1	2	3	4	5
Mit wem spricht du Englisch? *(Bitte keine Namen: Freund, Freundin, Verwandte…)* ➡ _____					
auf Englisch denken / mit dir selbst reden	1	2	3	4	5
englische Wörter in anderen Sprachen verwenden	1	2	3	4	5

My favourite sport

Sport are a favourite hobby for most people. Sports are played and watched on TV all over the world. What is your favourite sport and what can you tell me about it?

Write a text describing your favourite sport. You should write about:

- What sport is it?
- Where / when / how often do you play it?
- Why do you like it?

Write between **100-120 words**.

A family wants to spend the holidays at sea. They are driving there by car. While they are driving, they have some problems.

Figure 1 AI-generated picture using gencraft

In your **short story** you should:

- Tell the reader where they want to go
- Explain what happens on the way there
- Describe what happens in the end

Give your story a **title**. Write between **100 – 120 words**.

Summary variable	Individual variable
listening to music	listening to music on the phone
	listening to music on music apps
watching movies	watching movies on TV
	watching movies on the internet with subtitles
	watching movies on the internet without subtitles
	watching movies at the cinema
watching series	watching series on TV
	watching series on the internet with subtitles
	watching series on the internet without subtitles
playing video games	playing video games on the tablet or phone
	playing video games on PC or laptop
	playing multiplayer online games
reading on social media	reading messages on social media
	reading comments or status updates on social media
writing on social media	writing messages on social media
	writing comments or status updates on social media

Abstract

The global presence of English in areas such as media, technology, and popular culture has significantly reduced the distance between learners of English and their target language. Students come into contact with English outside instructed school lessons on a more regular basis and time spend on English language activities outside school has surpassed the time in instructed settings, due to the easy access to the internet and digital media. In research, different terminology for language learning in this context has been used. This study uses 'extramural English' (EE) to refer to English students engage in outside the walls of the classroom. The aim of this thesis is to examine what EE activities participants engage in, how often they come into contact with EE, and how EE correlates with lexical diversity in writing.

The participants in this master thesis are students from a second grade at a middle school in Vienna's 10th district. Data collection was conducted using a questionnaire that gathered information on the learners' demographic and linguistic backgrounds, their access to the internet and digital media, general leisure activities, and activities related to extramural English. Additionally, each student produced two written texts, which were used as the basis for calculating lexical diversity.

The results show similar results to previous studies conducted in this field. Younger learners engage less frequently in EE than older students. Overall, the most popular activities are playing video games, listening to music, and watching movies or series. Learners mostly favour receptive activities over productive ones. While male participants play video games more frequently, female participant favour listening to music. There were no statistically significant differences between genders in terms of frequency of engagement. Least popular activities are writing and reading. Furthermore, the results demonstrate that frequent EE engagement positively correlates with LD scores. While quantity of EE shows higher positive correlations with LD than most EE activities, the biggest effects can be observed for niche activities only few students engage in, such as writing, reading, speaking, or self-talk.

Die globale Präsenz der englischen Sprache in Bereichen wie Medien, Technologie und Pop-Kultur hat die Distanz zwischen Lernenden und der Fremdsprache English deutlich verringert. Dadurch kommen Schüler*innen häufiger in Kontakt mit authentischem Englisch außerhalb des Regelunterrichts. Der einfachere Zugang zum Internet und digitalen Medien hat dazu geführt, dass Jugendliche mehr Zeit mit Englisch außerhalb der Schule verbringen als innerhalb. Sprachen lernen in diesem Kontext wird in der Wissenschaft unter verschiedenen Begriffen behandelt. Diese Studie verwendet ‚extramural English' als Begriff für Englisch mit dem Lernende außerhalb der Wände des Klassenzimmers in Kontakt kommen. Das Ziel dieser Arbeit ist es zu ergründen, wie der Kontakt mit Englisch außerhalb der Schule geartet ist und wie dieser sich auf die Vielfalt des benutzen Vokabular in Texten auswirkt.

Die Teilnehmer*innen dieser Masterarbeit sind Lernende einer zweiten Klasse in einer Mittelschule im 10. Wiener Gemeindebezirk. Die Datenerhebung wurde mittels eines Fragebogens über demographischen und sprachlichen Hintergrund der Lernenden, deren Zugang zum Internet und digitalen Median, generelle Freizeitaktivitäten, sowie Aktivitäten im Zusammenhang mit ‚extramural Englisch' durchgeführt. Außerdem wurden von allen Lernenden je zwei Texte verfasst, die als Grundlage zur Berechnung der lexikalischen Vielfalt dienen.

Die Ergebnisse zeigen Gemeinsamkeiten mit vorherigen Studies in diesem Feld. Die jüngeren Schüler*innen dieser Studie kommen weniger häufig in Kontakt mit Englisch als ältere Schüler*innen. Die Lernenden bevorzugen rezeptive Aktivitäten über produktiven. Die beliebtesten Freizeitbeschäftigungen sind Videospiele spielen, Musik hören und Filme oder Serien schauen, wobei männliche Teilnehmer mehr Videospiele spielen und weibliche Teilnehmerinnen mehr englische Musik hören. Bei der Häufigkeit des Kontaktes mit Englisch gibt es jedoch im Unterschied zu vorherigen Studien keine geschlechterspezifische Unterschiede. Unbeliebt sind Aktivitäten wie Schreiben und Lesen. Die Ergebnisse zeigen auch, dass sich regelmäßiger Kontakt mit Englisch außerhalb der Schule positiv auf die lexikalische Vielfalt auswirkt. Während die reine Quantität an Kontakt mit Englisch meist höhere Effekte als die meisten Aktivitäten aufweist, können die größten Effekte bei Aktivitäten festgestellt werden, die nur wenige Lernende in ihrer Freizeit machen. Diese sind Schreiben, mit anderen Personen persönlich sprechen und Selbstgespräche oder Denken in Englisch.