

DICTIONARIES, DEEPL, AND CHATGPT: COMPARING LANGUAGE TOOL EFFECTS ON L2 ENGLISH SPEAKING AND LEARNER PERCEPTIONS

A MULTI-METHOD EXPERIMENTAL STUDY

Word count: 18.312

Irma Schampheleer

Student number: 02111525

Supervisors: Prof. Dr. Lieve Macken, Prof. Dr. Vanessa De Wilde

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of
Master of Arts in Interpreting (Dutch, French, and English)

Academic year: 2024-2025

Abstract

In recent years, language learners have increasingly relied on digital tools such as machine translation (MT) systems and generative artificial intelligence (AI) tools for language support. While the effects of MT tools like DeepL on second language (L2) writing have been relatively well-documented, fewer studies have examined their influence on L2 speaking performance. Research on the impact of generative AI tools such as ChatGPT in this context is even more limited. This multi-method experimental study aims to assess how the use of MT and generative AI tools affects L2 English speaking performance, as well as learners' experiences preparing oral presentations. Thirty university language students were divided into three groups and assigned either dictionaries, DeepL, or ChatGPT for language support when preparing a group presentation. The presentations were recorded, transcribed, and analysed for fluency and lexical complexity. Additionally, participants completed a two-part survey: one part reflecting on their assigned tool during the preparation process, and one evaluating their general experience with all three tools. Results showed that MT users outperformed other groups in fluency-related measures, while AI users demonstrated greater lexical diversity and density. Lecturer evaluations favoured presentations prepared with MT and AI tools over those prepared with dictionaries. Additionally, participants using notes during the delivery demonstrated a higher lexical diversity and density. Survey responses indicated that perceived enjoyment and usefulness varied between groups, with generative AI tools generally rated more positively for idea generation but less so for practical usefulness. Overall, participants favoured MT tools for preparing both oral and written tasks. These findings suggest that while MT tools currently offer the most consistent benefits for L2 speaking performance, generative AI tools hold promise, particularly for idea development and lexical variation. (281 words)

Table of contents

1	Introduction	9
2	Literature review	11
2.1	<i>L2 production</i>	<i>11</i>
2.1.1	L2 fluency	11
2.1.2	L2 complexity	12
2.1.3	L2 accuracy	13
2.1.4	L2 pronunciation	14
2.2	<i>Impact of MT use on L2 production.....</i>	<i>14</i>
2.2.1	Methodological approaches of previous studies on MT use for L2 writing	14
2.2.2	Methodological approaches of previous studies on MT use for L2 speaking.....	19
2.2.3	Main findings of studies on the impact of MT use on L2 writing	23
2.2.4	Main findings of studies on the impact of MT use on L2 speaking.....	26
2.3	<i>Impact of generative AI use on L2 production.....</i>	<i>27</i>
2.4	<i>Research questions.....</i>	<i>29</i>
3	Methodology	33
3.1	<i>Participants.....</i>	<i>33</i>
3.2	<i>Speaking task.....</i>	<i>33</i>
3.3	<i>Language tools.....</i>	<i>35</i>
3.4	<i>Recordings and transcriptions</i>	<i>37</i>
3.5	<i>Survey.....</i>	<i>37</i>
3.6	<i>Analysis</i>	<i>38</i>
4	Speaking performance	39
4.1	<i>Fluency.....</i>	<i>39</i>
4.2	<i>Lecturer evaluation.....</i>	<i>40</i>
4.3	<i>Lexical complexity.....</i>	<i>41</i>
4.4	<i>Presentation aid use.....</i>	<i>43</i>
5	Preparation process and reflections on tool usage	45
5.1	<i>Preparation time</i>	<i>45</i>
5.2	<i>Language tool use</i>	<i>46</i>
5.2.1	DI-group	46
5.2.2	MT-group	48
5.2.3	AI-group.....	50
5.3	<i>Attitudes towards assigned tool</i>	<i>54</i>
5.3.1	Perceived enjoyment	54
5.3.2	Perceived usefulness	55

6	General experience with language tools	57
6.1	<i>Language tool use</i>	57
6.1.1	Dictionaries	57
6.1.2	Machine translation.....	58
6.1.3	Generative AI	60
6.2	<i>Attitudes towards language tools.....</i>	63
6.2.1	Perceived enjoyment	63
6.2.2	Perceived usefulness	65
6.2.3	Perceived advantages and disadvantages	66
6.2.4	Preference.....	68
7	Discussion	71
8	Conclusion	75
	Appendix	79
A.	<i>Assignment sheets</i>	79
B.	<i>Survey questions</i>	82
C.	<i>Prompts entered into ChatGPT by the AI-group</i>	91
D.	<i>Perceived advantages and disadvantages of dictionaries, machine translation, and generative AI tools</i> 94	
E.	<i>Transcriptions.....</i>	99

List of abbreviations

Abbreviation	Meaning
AI	artificial intelligence
AI-group	group of participants in this study that was instructed to exclusively use ChatGPT for language support during the preparation stage of their presentations
ASR	automatic speech recognition
AWL	Academic Word List
DI-group	group of participants in this study that was instructed to exclusively use (online) dictionaries for language support during the preparation stage of their presentations
EFL	English as a foreign language
L1	first language
L2	second language
MT	machine translation
MT-group	group of participants in this study that was instructed to exclusively use DeepL for language support during the preparation stage of their presentations
TTS	text-to-speech

List of tables and figures

Table 1: Comparison of five studies investigating the impact of MT use on L2 writing in terms of their aim, their participants' learner profiles, the writing tasks the studies analysed, and their method of data analysis	15
Table 2: Comparison of four studies investigating the impact of MT use on L2 speaking in terms of their aim, their participants' learner profiles, the content of the speaking exercise the studies contain, and their method of data collection	20
Table 3: Overview of the numbers of students in each language tool group, and the topics they selected for their group presentations	36
Table 4: Overview of the average word count, duration, and speech rate (words per minute) for each language tool group	39
Table 5: Average presentation fluency, marks, and lexical complexity depending on the use of slides	43
Table 6: Average presentation fluency, marks, and lexical complexity depending on the use of notes	44
Table 7: Examples of prompts by participants in the AI-group divided into three categories	50
Table 8: Summary of the perceived advantages and disadvantages of dictionaries, MT tools, and generative AI tools	67
Figure 1: Simplified visualisation of the components and subcomponents of L2 complexity, inspired by Housen et al. (2012). On the right side, examples of measures used to quantify the subcomponents are provided.	13
Figure 2: Speech rate (words/min) dispersion per language tool group	40
Figure 3: Mark dispersion per language tool group	41
Figure 4: Type-token ratio and lexical density dispersion across the three language tool groups	42
Figure 5: Average vocabulary profile per language tool group	43
Figure 6: Preparation time (min) dispersion per language tool group	45
Figure 7: Overview of the dictionaries participants in the DI-group reported having used to prepare their presentations (total number of participants: 9)	47
Figure 8: Overview of the types of information participants in the DI-group reported looking up in dictionaries for their presentations (total number of participants: 9)	47
Figure 9: Type of input entered into DeepL by the MT-group (total number of participants: 7)	48
Figure 10: Ways in which DeepL was beneficial during the preparation of the presentation according to the MT-group (total number of participants: 7)	49
Figure 11: Acceptance rate of the DeepL output by the MT-group (total number of participants: 7)	49
Figure 12: Ways in which ChatGPT was beneficial during the preparation of the presentation according to the AI-group (total number of participants: 14)	51
Figure 13: Stage at which participants in the AI-group found ChatGPT to be especially helpful (total number of participants: 14)	51

Figure 14: Acceptance rate of the ChatGPT output by the AI-group (total number of participants: 14).....	52
Figure 15: Reasons why the AI-group did not accept the ChatGPT output (total number of participants: 14).....	52
Figure 16: Stage at which participants in the MT-group and AI-group found their assigned tool to be especially beneficial	53
Figure 17: Acceptance rates of the DeepL and ChatGPT output by the MT-group and AI-group, respectively.....	54
Figure 18: Perceived enjoyment and usefulness dispersion across the three language tool groups	55
Figure 19: Overview of the dictionaries participants reported having worked with before (total number of participants: 30)	57
Figure 20: Overview of the types of information participants reported usually looking up in dictionaries (total number of participants: 30)	58
Figure 21: Overview of the machine translation tools participants reported having worked with before (total number of participants: 30).....	59
Figure 22: Overview of the functional purposes for which participants use MT tools (total number of participants: 30)	60
Figure 23: Comparison of MT and generative AI tools in terms of the participants' use for the four language pillars (total number of participants: 30 for MT use, 29 regarding the use of generative AI).....	61
Figure 24: Overview of the functional purposes for which participants use generative AI tools (total number of participants: 29)	62
Figure 25: Comparison of MT and generative AI tools in terms of the participants' motivation to use them (total number of participants: 30 for MT use, 29 regarding the use of generative AI).....	63
Figure 26: Overview of the general perceived enjoyment of the three types of language tools for language learning, ranked on a scale of 1 to 5 stars.....	64
Figure 27: Comparison of the perceived enjoyment of the three types of language tools for language learning across the three groups, ranked on a scale of 1 to 5 stars	64
Figure 28: Overview of the general perceived usefulness of the three types of language tools for language learning, ranked on a scale of 1 to 5 stars.....	65
Figure 29: Comparison of the perceived usefulness of the three types of language tools for language learning across the three groups, ranked on a scale of 1 to 5 stars	66
Figure 30: Distribution of the preferred language tool for preparing oral tasks in a foreign language, ranked on a scale of 1 to 3 (total number of participants: 30).....	68
Figure 31: Distribution of the preferred language tool for preparing written tasks in a foreign language, ranked on a scale of 1 to 3 (total number of participants: 30).....	69

PREAMBLE

Artificial intelligence tools were employed to support the writing process of this paper. Both DeepL Write and ChatGPT were used alternately for language assistance across all chapters.

PREFACE

This thesis marks the final project of my master's degree and gave me the opportunity to bring together two fields that have always interested me: second language acquisition and the role of technology in education. Throughout my studies, I have been fascinated by how we acquire a second language and how digital tools can support that process. During my master in Interpreting this year, with its strong focus on spoken communication, I was able to connect these interests and explore them in the context of L2 speaking.

I would like to thank my thesis supervisors for their valuable feedback, both during online and in-person meetings. Their thoughtful comments not only helped refine my research but also encouraged me to keep sight of the broader perspective when I became too caught up in the finer details.

I am also grateful to my bachelor's thesis supervisor, Margot Fonteyne, whose insights into this topic helped spark my interest and who taught me a great deal about structuring academic writing, building clear arguments, and presenting research findings effectively.

I am especially grateful to Mr. Chan for welcoming me into his speaking classes and for allowing me to record his students' speaking skills. His cooperation made it possible for me to collect authentic student presentations, which was invaluable for my study.

To all the students who participated in this research: thank you for your time, enthusiasm, and well-prepared presentations. I was genuinely impressed by your level of English and the thoughtful arguments you presented.

Finally, I want to thank my friends and family who have cheered me on throughout this journey. Their encouragement, patience, and advice, especially during the final months of writing, meant more than they could imagine.

This project has allowed me to combine two areas of study that are important to me, and I hope it contributes to ongoing discussions about the role of technology in language education.

1 INTRODUCTION

In recent years, digital language tools such as machine translation systems and generative AI models have become increasingly integrated into language learning environments. Tools like DeepL and ChatGPT are now readily accessible and widely used by language learners across all proficiency levels. Their growing popularity raises important questions about how such technologies influence L2 learning processes and outcomes, particularly in productive skills like speaking.

While a substantial body of research has explored the use of machine translation tools for L2 writing, their impact on speaking performance remains less studied. Even fewer studies have experimentally investigated the role of generative AI tools such as ChatGPT in L2 speaking and writing. As these technologies evolve and become more sophisticated, it becomes increasingly relevant to evaluate not only how learners use them, but also how these tools impact learners' linguistic output and shape their learning experience.

This study aims to address that gap by examining the effects of three types of digital language tools – dictionaries, machine translation (DeepL), and generative AI (ChatGPT) – on L2 English speaking performance. To that end, we conducted a multi-method experimental study involving thirty university language students. The participants were divided into three groups and invited to prepare and deliver group presentations using one of the tools as their main source of linguistic support. Their presentations were recorded, transcribed, and analysed in terms of fluency, lexical complexity, lecturer evaluation, and contingent on the use of presentation aids. In addition, a two-part survey was administered to gather insights into the students' experiences using their assigned tools during the preparation process, as well as their general attitudes and preferences regarding all three tools.

The aim of this paper is twofold. First, it seeks to compare the linguistic outcomes of students who prepared their presentations with different tools (i.e. dictionaries, machine translation, and generative AI), in order to evaluate how each tool may impact spoken production in English as a second language. Second, it aims to better understand students' motivations for using these tools, the ways in which they integrated them into their preparation process, and their perceived usefulness and enjoyment. Through this combined lens of performance and perception, the study contributes to the growing field of research on technology-enhanced language learning

and offers insights into the pedagogical implications of integrating MT and AI tools in communicative language tasks.

2 LITERATURE REVIEW

This chapter provides an overview of the theoretical and empirical foundations relevant to this study. It begins by outlining the main aspects of L2 production (2.1), including fluency, complexity, accuracy, and pronunciation, which serve as the primary dimensions for evaluating L2 speaking performance. The chapter then reviews existing research on the impact of MT use on L2 production (2.2), focusing on both writing and speaking tasks. Particular attention is given to the methodological approaches employed in previous studies and their main findings. Following this, recent studies on the use of generative AI tools in L2 production are discussed (2.3), as this emerging area is rapidly gaining attention in language education. The chapter concludes by presenting the research questions guiding the present study (2.4).

2.1 L2 production

This section delves into the theoretical foundations that are essential for a nuanced understanding of the impact of language tool use on L2 speaking. In the context of L2 production, three main components are frequently addressed in research: *fluency*, *complexity*, and *accuracy*. Skehan (1998) was the first to bring these three components together in a proficiency model. The dimensions serve as critical indicators of language proficiency (Housen et al., 2012) and are essential for evaluating the effectiveness of technology integration in language learning. Furthermore, as most studies in the literature review investigating the impact of MT use on L2 speaking focus on pronunciation, this dimension is also discussed. This section provides a concise overview of the four aspects, examining their definitions, significance, and their operationalisations in empirical studies.

2.1.1 L2 fluency

A first component of L2 production is fluency, “a learner’s or user’s global language proficiency, particularly as characterized in terms of the ease, eloquence, ‘smoothness’ and native-likeness of speech or writing” (Housen et al., 2012, p. 4). Furthermore, it is argued that the concept of fluency itself is multidimensional, too. *Speed fluency* (i.e. the rate of delivery), *breakdown fluency* (i.e. pausing behaviour), and *repair fluency* (i.e. the frequency of repetitions and self-corrections) are three of the subdimensions of fluency (Housen et al.; Lambert & Kormos, 2014). In light of this definition, fluency can be considered a phonological phenomenon, in contrast to accuracy and complexity. Nevertheless, fluency is equally pertinent

when it comes to writing. Not only are brief pauses and few revisions typical of fluent writing processes, so is a high production rate (Van Waes & Leijten, 2015). This last subcomponent is particularly frequently measured in writing studies by counting the total number of words or by calculating the number of words produced per minute (Kellogg, 1996). The latter is currently the most widely recognised measure of writing fluency (Van Waes & Leijten).

2.1.2 L2 complexity

The concept of L2 complexity is often defined as “the ability to use a wide and varied range of sophisticated structures and vocabulary in the L2” (Housen et al., 2012, p. 2). This working definition was developed in the 1990s and is still used. However, according to Housen et al., the construct of linguistic complexity is more complex than previously thought. A simplified visualisation of their taxonomy is given below (

Figure 1).

In particular, linguistic complexity consists of two major dimensions: *grammatical* and *lexical complexity*. The first dimension, grammatical complexity, can be divided into *syntactic* and *morphological complexity*. However, for the purposes of this paper, it is necessary to identify quantifiable measures. For this variable, the more finely-grained subcomponents of syntactic and morphological complexity, such as subclause ratios (a measure often referred to as *subordination*), length measures (e.g. *words per T-unit*, or *words per clause*), and the number of verb forms, are of particular relevance. Lexical complexity, then, “a learner’s depth of lexical knowledge” (Bulté et al., 2008), is often divided into three dimensions: *density*, *diversity*, and *sophistication* (Housen et al., 2012). Examples of measures used to quantify these variables include *lexical words / function words* and *lexical words / total words* for lexical density, *number of word types* and *type-token ratio* for diversity and *less frequent words / total words* for sophistication.

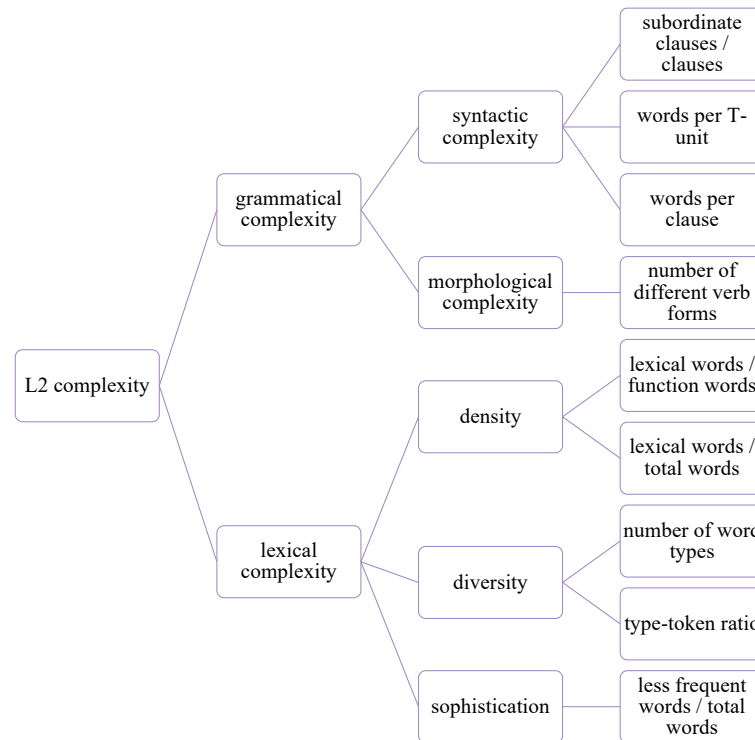


Figure 1: Simplified visualisation of the components and subcomponents of L2 complexity, inspired by Housen et al. (2012). On the right side, examples of measures used to quantify the subcomponents are provided.

2.1.3 L2 accuracy

Accuracy in L2 production is generally characterised as “the ability to produce target-like and error-free language” (Housen et al., 2012). In other words, this concept refers to the degree to which a learner’s L2 production differs from a set norm, typically that of a native speaker (Housen et al.). Such deviations are often referred to as ‘errors’. However, depending on the norm, such deviations can simply be less appropriate or acceptable in the L2 rather than erroneous. Consequently, Housen et al. propose that accuracy should also encompass appropriateness and acceptability.

As outlined by Lambert and Kormos (2014), the accuracy of grammatical structures is commonly evaluated through two methodologies. The first approach entails computing the ratio of errors present in a text to a predefined unit of production (e.g., words, clauses, sentential units). The second approach involves calculating the proportion of these units that remain error-free.

2.1.4 L2 pronunciation

L2 pronunciation is a notoriously challenging dimension to quantify in research. Traditionally, a native-like pronunciation is considered the ultimate goal, and therefore L2 pronunciation is evaluated using a native model as a reference point. However, this norm has been reconsidered in recent years, with *intelligibility* and *comprehensibility* now being common constructs to evaluate pronunciation. They both refer to the extent to which L2 speech is understandable. Intelligibility is defined as the proportion of L2 speech understood by listeners. This construct is frequently operationalised through transcription tasks. Comprehensibility, then, refers to the amount of effort required to understand the L2 speech, operationalised via rating scales (Mairano et al., 2023).

2.2 Impact of MT use on L2 production

In this section, previous studies on the impact of MT use on L2 production are discussed. Section 2.2.1 provides an overview of the methodologies in studies on the impact of MT-assisted L2 writing. In Section 2.2.2, you will find a similar discussion of studies measuring the impact of MT use on L2 speaking. Sections 2.2.3 and 2.2.4 subsequently discuss the results of these studies.

2.2.1 Methodological approaches of previous studies on MT use for L2 writing

To gain insight into language learners' MT use and its impact on their L2 writing, it is crucial to examine how this topic has been researched in the past. Therefore, this section will provide an overview of the methodologies used in a number of studies that have investigated how the use of MT affects writing in a foreign language. We will focus on the relevant objectives of these studies, the learner profiles of their participants, the writing tasks in which the participants engaged, and lastly, the methods of data analysis adopted. Table 1 summarises five studies in terms of these characteristics.

	Chung and Ahn (2022)	Kol et al. (2018)	Lee (2020)	Stapleton (2021)	Tsai (2019)
Relevant objective(s)	Determine how MT use affects CALF ¹ measures in L2 writing across different proficiency levels and text genres; investigate student perceptions of MT use	Determine whether MT use affects the quality and quantity of advanced L2 writing; investigate student perceptions of MT use	Investigate how MT use while editing improves students' L2 writing and their revision strategies; investigate student perceptions of MT use	Compare machine-translated texts to parallel texts directly written in the L2 by the same individuals	Compare machine-translated texts to parallel texts directly written in the L2 by the same individuals; investigate student perceptions of MT use
Number of participants	91	25	34	26	124
Participants' age	Not specified, first-year college students	Not specified, university students	Not specified, university students	11	Not specified, university students
Participants' L1 and L2	Korean; English	Hebrew; English	Korean; English	Chinese; English	Chinese; English
Participants' proficiency level in L2	Not mentioned	B2	Intermediate to high-intermediate	Not mentioned	B1-B2
Non-MT text	Self-written narrative or argumentative text (depending on group)	Self-written L2 text (dictionary allowed)	Translation of a self-written L1 text into L2	Self-written L2 text	Self-written L2 text
MT text	Self-written narrative or argumentative text (depending on group) with MT help	Self-written L2 text with MT help	Revised self-translated L2 text (revisions based on MT version; other resources allowed)	MT of a parallel L1 text	Revised self-written L2 text (revisions based on MT of parallel L1 text)
Time allotted for each writing task	Not mentioned	40 minutes	Not mentioned	60 minutes	30 minutes
Data analysis	Quantitative computational and human assessment: Syntactic and Lexical Complexity Analyzer, error-free clause ratio and weighted clause ratio, number of words / time on task	Quantitative computational and human assessment: writing score, number of words, vocabulary profile using VocabProfiler	Quantitative human assessment: overall quality, lexical and grammatical errors, lexical and syntactic complexity	Qualitative human assessment: patterns of improvement separated into three categories (grammar, vocabulary, and cognitive load)	Quantitative computational assessment using 1Checker and VocabProfiler

Table 1: Comparison of five studies investigating the impact of MT use on L2 writing in terms of their aim, their participants' learner profiles, the writing tasks the studies analysed, and their method of data analysis

¹ Syntactic complexity, accuracy, lexical complexity, and fluency

2.2.1.1 Objective

Studies in this field often have different objectives. Some aim to compare MT-assisted writing to texts directly written in students' L2, either with (Kol et al., 2018) or without (Chung & Ahn, 2022) other resources such as dictionaries. Others seek to investigate the difference between self-written or -translated L2 texts and automatically translated versions of L1 texts written by the same language learners (Lee, 2020; Tsai, 2019; Stapleton, 2021).

These studies also differ in terms of the features they aim to measure. Chung and Ahn (2022), for instance, explored how complexity, accuracy, and fluency are influenced when learners are allowed to use MT while writing. Accuracy is an especially commonly researched measure, as seen in the works of Kol et al. (2018), Lee (2020), Stapleton (2021), and Tsai (2019), who investigated, among other things, whether MT use reduced language errors. Vocabulary has also frequently been evaluated – both Kol et al. and Tsai were interested in the vocabulary profile of student writing before and after MT use, i.e. changes in the percentage of common, less common, and academic words in their texts; Lee (2020) aimed to investigate differences in lexical complexity (lexical density and diversity) between self- and machine-translated L2 texts. In addition, numerous studies aim to examine whether the use of MT increases the length of the texts produced by language learners (Kol et al., Tsai). One study in this overview also measured the effect of two moderator variables, namely text genre and proficiency level (Chung & Ahn).

In addition to examining the effects of MT use on L2 writing outcomes, several studies have focused on learners' perceptions. Chung and Ahn (2022) aimed to explore how students perceive and use MT in English composition. Similarly, Kol et al. (2018) and Lee (2020) included objectives related to understanding students' attitudes towards MT use. Tsai (2019) specifically investigated whether learners considered Google Translate a useful tool for improving aspects of EFL writing, such as vocabulary, grammar, and writing speed, as well as their willingness to continue using it in the future.

2.2.1.2 Learner profiles

In order to assess the generalisability of research findings on the effect of MT use in L2 writing, it is also important to examine the characteristics of the study participants. Firstly, it is clear that the participants in these studies are often university students (Chung & Ahn, 2022; Kol et

al, 2018; Lee, 2020; Tsai, 2019). One study in this overview, however, involved 11-year-old primary school students (Stapleton, 2021). Additionally, all participants in the reviewed studies were EFL learners, despite having different first languages, such as Korean (Chung & Ahn; Lee), Chinese (Stapleton; Tsai) or Hebrew (Kol et al.). In the majority of the studies, participants were reported to have an intermediate level of English at the time of the study (Kol et al., Lee, Tsai).

2.2.1.3 Data collection

In each reviewed study, participants were instructed to produce two texts. In some studies, they initially wrote a text without any MT help, and subsequently with its assistance, both in their second language (Chung & Ahn, 2022; Kol et al., 2018). Other studies required participants to write two texts on the same topic, one in their first language and one in their second language. The L1 text was then fed into an automatic translator, most often Google Translate (Stapleton, 2021; Tsai, 2019). Lee (2020) adopted a unique approach – participants first translated their one-page self-written texts from their first to their second language without the aid of MT, but with the use of other resources. They then used MT to translate the same self-written L1 text into English and finally edited their self-translated text with the help of their machine-translated version and other resources.

The writing tasks frequently revolved around school-related topics. Chung and Ahn (2022) conducted a study in which participants were requested to write about either their first day of college or their last day of high school, resulting in narrative texts. The argumentative texts in the same study focused on school uniforms in high school or the use of cell phones in class. In another study (Stapleton, 2021), students were asked to respond to the question “Is a half-day school a good idea?” The topic in Lee’s (2020) study was language-related – participants were instructed to write a text about the language young people use to text today. Tsai (2019) assigned different topics based on the student’s level. Sophomore students were tasked with writing an essay on the significance of participating in international trade fairs for local trade companies in Taiwan. Junior students were required to compare a line and a pie graph in a marketing context, while senior students were asked to reflect on the timing and fairness of applying justice after watching a 5-minute film extract. In Kol et al.’s (2018) study, participants provided their views on a topic that had been discussed during the course in which they were enrolled.

Participants were given 30 minutes (Tsai, 2019) to an hour (Stapleton, 2021) to complete each task.

Additionally, several studies on MT use in L2 writing have gathered learners' perceptions through post-task surveys, questionnaires, and interviews. Chung and Ahn (2022) and Kol et al. (2018) used short surveys or questionnaires after the second writing task to capture students' experiences with MT. Lee (2020) combined interviews with reflection papers to explore learners' perceptions and revision intentions. Tsai (2019) collected feedback using a survey with both Likert-scale items and open-ended questions. These methods provided insight into how learners perceived MT's role in their writing process.

2.2.1.4 Data analysis

Most studies in this overview took a quantitative approach to measure the impact of MT use on L2 writing quality (Chung & Ahn, 2022; Kol et al., 2018; Lee, 2020; Tsai, 2019). Furthermore, many research designs included human raters (Chung & Ahn; Kol et al; Lee; Stapleton, 2021). When two assessors were involved, interrater reliability was calculated (Chung & Ahn, Lee). Some studies opted for computational assessment of the student writing output (Tsai), sometimes in addition to human evaluation (Kol et al.).

In studies that employed a grammatical complexity measure, the focus was on syntactic complexity. Chung and Ahn (2022) used 12 of the 14 measures from Lu's (2010) L2 Syntactic Complexity Analyser to gain insight into any differences in unit lengths (i.e. mean lengths of sentences, clauses, and T-units) and frequency of structures with subordination, coordination, complex nominals, and verb phrases. Lee (2020) also calculated syntactic complexity, although the methodology employed was not specified.

Lexical complexity, a measure examined by all studies included in this overview, was also analysed in various ways. Chung and Ahn (2022), for instance, utilised the Lexical Complexity Analyser designed by Lu (2012), which analyses 15 lexical complexity measures, to record any differences in lexical density, diversity, and sophistication. Both Tsai (2019) and Kol et al. (2018) employed the VocabProfiler by Lextutor (www.lextutor.ca/vp/eng/) to analyse the so-called vocabulary profile of the participants' self-written and MT-aided texts. The online tool enabled the researchers to compare the percentages of low-frequency and high-frequency vocabulary of the two types of texts. In addition, Tsai calculated the lexical density of the texts

in order to gain insight into the amount of information they convey. Additionally, Lee (2020) calculated lexical complexity, encompassing lexical diversity and density, although the specific methodology was not specified. As previously mentioned, lexical errors were also assessed in Lee's study. Lastly, Stapleton (2021) qualitatively analysed patterns of improvement in terms of vocabulary after MT use.

To measure accuracy (sometimes referred to as 'writing score' or 'writing quality'), one study in this overview simply counted the number of errors (Lee, 2020). Most studies, however, considered the percentage or probability of writing mistakes or error-free clauses in a text (Chung & Ahn, 2022; Kol et al., 2018; Tsai, 2019). Furthermore, the errors included in this measure were predominantly grammatical (Chung & Ahn; Kol et al.; Lee; Tsai); Lee also included incorrect word choice in a given context, and Tsai took spelling errors into account in his accuracy measure.

Almost all studies in this overview examined differences in length (also described as 'writing quantity' or 'fluency') before and after MT use by counting the number of words (Chung & Ahn, 2022; Kol et al., 2018; Tsai, 2019). Stapleton (2021) used text length differences to gain insight into the effects of MT use on cognitive load.

For the perception-related data, Chung and Ahn (2022) analysed survey responses descriptively by calculating percentages for closed questions and conducting thematic analysis on open-ended responses. Lee (2020) used multiple rounds of coding to identify themes in interviews and reflection papers, which were then triangulated with the writing outcomes. Tsai (2019) processed questionnaire data quantitatively using SPSS.

2.2.2 Methodological approaches of previous studies on MT use for L2 speaking

This section focuses on how researchers have recorded the effect of MT integration in speaking activities. We will first examine the different aims of these studies, before discussing the learner profiles of their participants. Thirdly, we will elaborate on the content of the speaking exercises used, and finally, we will explore the different data collection methods employed. Table 2 briefly compares four studies in terms of these features.

	He and Cardoso (2021)	Khademi and Cardoso (2022)	Rushton (2022)	Sakamoto (2022)
Relevant objective(s)	Determine whether English learners can acquire the pronunciation of the past <i>-ed</i> allomorph in terms of phonological awareness, phonemic discrimination, and/or oral production using MT but without guidance	Determine whether English learners can acquire the pronunciation of the past <i>-ed</i> allomorph in terms of awareness, perception, and/or production using MT	Motivate students to produce longer speaking turns using MT	Qualitatively analyse students' perceptions of MT used for a presentation activity
Number of participants	29	20	36	111
Participants' age	16-18	30-40	not specified, university students	not specified, university students
Participants' L1 and L2	Chinese – English	Farsi – English	Japanese – English	Japanese – English
Participants' proficiency level in L2	Intermediate	Intermediate	A1-B1	A2-B1
Speaking exercise	Read aloud; engage in a role-play	Read out a word list with past <i>-ed</i> forms; engage in a role-play interview	Make a draft with MT; talk for 30 seconds to 2 minutes	Make a draft with MT; give a 3-minute oral presentation about a subject of choice
Data collection	Background questionnaire; pre- and post-test to measure improvements in the participants' pronunciation of the past <i>-ed</i> allomorph	Pre- and post-test to measure improvements in the participants' pronunciation of the past <i>-ed</i> allomorph	Survey to assess students' perceptions	Survey to assess students' perceptions

Table 2: Comparison of four studies investigating the impact of MT use on L2 speaking in terms of their aim, their participants' learner profiles, the content of the speaking exercise the studies contain, and their method of data collection

2.2.2.1 Objective

Studies examining how MT use affects the quality of spoken output amongst language learners can have different purposes. Khademi and Cardoso (2022), for example, investigated the use of MT to support L2 learners in the acquisition of a particular aspect of English pronunciation, with a focus on the learners' awareness, perception and production. More specifically, they aimed to examine whether L2 learners of English could learn how to pronounce the past *-ed* allomorph through Google Translate's speech technologies (both automatic speech recognition (ASR) and text-to-speech (TTS)). He and Cardoso (2021) adopted a similar approach – they aimed to examine whether EFL learners can acquire the pronunciation of the same allomorph, albeit through Microsoft Translate's TTS and ASR functionalities. The features studied also differed from Khademi and Cardoso's research design – in this study, the first three stages of the communicative framework for teaching English pronunciation by Celce-Murcia and her colleagues (2010) were examined, namely phonological awareness (or 'description and analysis', as mentioned in the original framework), phonemic (or 'listening') discrimination, and oral production (or 'controlled, guided, and communicative practice'). Rushton (2022), on the other hand, focused, among other things, on motivating Japanese students to produce longer turns when speaking English. Lastly, Sakamoto (2022) aimed to evaluate students' perceptions of MT when used for an oral presentation activity.

2.2.2.2 Learner profiles

It is also interesting to compare the learner profiles of the participants in these studies in terms of their age, first and second language, and proficiency level. As shown in Table 1, the participants engaging in Rushton's (2022) and Sakamoto's (2022) studies had similar learner profiles. In both cases, the participants were Japanese-speaking university students learning English as a second language. At the time of the research, they had an A1-B1 or A2-B1 CEFR level for English. Khademi and Cardoso (2022) also conducted a study involving learners of English as a foreign language. Their first language, however, was Farsi. They were aged between 30 and 40 and were intermediate speakers of English. Lastly, the participants engaging in He and Cardoso's study (2021) were also intermediate EFL students, but they were native speakers of Chinese and between the ages of 16 and 18.

These studies included varying numbers of participants, ranging from 29 (He & Cardoso, 2021) to 36 (Rushton, 2022), with the exception of Sakamoto's study (2022), which involved 111 students.

2.2.2.3 Speaking activity

Most studies quantifying the effect of MT use on oral production first require language learners to engage in a speaking activity. The exact assignment, however, often varies. Sakamoto (2022), for instance, conducted a study in which participants were given a presentation assignment. They were asked to briefly research a topic of their choice and then deliver a 3-minute oral presentation in English about their research. The students initially composed their texts in Japanese and were subsequently taught how to use DeepL. They then proceeded to draft their manuscripts in English using the MT tool. Similarly, Rushton (2022) asked participants to speak in English for 30 seconds to 2 minutes. The students were advised to use DeepL rather than other online tools, but no explicit instructions were provided in this study.

Both other studies discussed in this section (He & Cardoso, 2021; Khademi & Cardoso, 2022) opted for a pretest-posttest research design to measure students' awareness, perception or phonemic discrimination, and production of the English past *-ed* allomorph, with two tests for each of those features. Both studies first provided their participants with questions to assess their phonological knowledge. For aural discrimination, students were required to discriminate past from non-past constructions, and then identify the correct pronunciation of the *-ed* allomorph (/t/, /d/, or /id/). To measure improvements in production, participants in both studies completed two clearly defined speaking activities while being recorded. First, they were asked to read out a word list with past tense *-ed* forms. The second activity required them to engage in a role-play interview with one of the researchers. The treatment in both studies required the students to use either Microsoft Translate's (He & Cardoso) or Google Translate's (Khademi & Cardoso) TTS and ASR technologies to listen to and produce the target forms. After the treatment, they completed similar tests to the pre-tests, but with randomised items.

2.2.2.4 Data collection method

In terms of data collection, researchers have also employed different methods. Survey research appears to be the most commonly used method, for both Sakamoto (2022) and Rushton (2022)

opted for questionnaires in their studies. It is important to note that these surveys focused on the students' perceptions of MT use for the speaking activities they engaged in, rather than measuring the quality of the spoken output. He and Cardoso's (2021) and Khademi and Cardoso's (2022) research designs did, however, include post-tests to determine whether the participants' pronunciation of the *-ed* endings had improved after using MT. As previously mentioned, they measured improvements in terms of awareness, perception, and production.

2.2.3 Main findings of studies on the impact of MT use on L2 writing

Studies investigating the effect of MT use on writing in a second language, the methodologies of which we have discussed in section 2.2.1, have yielded interesting results regarding different linguistic features. In this section, we will explore the influence of MT assistance on a number of key linguistic dimensions. Throughout the sub-sections, we will also address genre-specific effects and proficiency level considerations where relevant, to provide a nuanced understanding of MT's role in different contexts.

2.2.3.1 Syntactic complexity

The impact of MT use on syntactic complexity in L2 writing reveals nuanced outcomes across studies. Chung and Ahn (2022) found that subordination scores significantly declined after MT assistance, with high-proficiency learners demonstrating higher scores compared to low-proficiency learners. Interestingly, low-proficiency learners exhibited increased coordinate phrase ratios post-MT. In contrast, high-proficiency learners experienced a decline in this measure. While argumentative writing demonstrated higher scores than narrative writing in several syntactic complexity measures related to clauses (e.g. clause length), verb phrases, complex nominals, and dependent clauses, narrative writing outperformed argumentative writing in terms of T-units². Conversely, Lee (2020) found no statistically meaningful differences in sentence complexity between the self-translated and edited MT-assisted versions of participants' texts. These findings suggest that while MT may influence certain aspects of syntactic complexity, its impact varies across proficiency levels, writing genres, and task types.

²A T-unit is defined as "a minimal unit constituting a complete sentence, consisting of one independent clause and any dependent clauses connected to it" (Collins Dictionary).

2.2.3.2 Lexical complexity

Research on the impact of MT use on lexical complexity in L2 writing, too, has yielded mixed results. Chung and Ahn (2022), for instance, highlight significant increases in lexical variation but note declines in lexical sophistication scores. While high-proficiency learners demonstrated higher lexical variation scores with MT assistance, their lexical sophistication scores decreased. Furthermore, high-proficiency writers' scripts in this study contained more word types than those written by the low-proficiency group. Kol et al. (2018) found that MT assistance resulted in improved vocabulary profiles, with students using fewer basic words and more advanced and academic vocabulary. Tsai (2019) also observed higher lexical density and advanced vocabulary in students' machine-translated texts compared to their self-written ones, suggesting a more accurate and sophisticated use of language with MT assistance. However, Lee (2020) found no significant changes in lexical diversity and density between the initial (self-translated) and final (edited) versions, indicating potential variability in the impact of MT on lexical aspects depending on the writing task the study entails. Overall, while MT can enhance the number of advanced and academic vocabulary items, its effects on lexical diversity, density, and sophistication may vary across different writing tasks and learner proficiency levels.

2.2.3.3 Accuracy

Studies that have examined differences in accuracy between MT-assisted writing and unaided L2 writing present varying outcomes. Chung and Ahn (2022), for instance, highlight significant improvements in accuracy, whereas Kol et al. (2018) found no significant difference between MT-assisted and unaided writing in terms of morphological and syntactical mistakes. Tsai (2019) observed fewer grammar and spelling mistakes in L1 automatically translated texts compared to self-written L2 texts, which indicates better grammatical accuracy with MT assistance. Additionally, Lee (2020) noted a significant decrease in total errors, lexical errors, and grammatical errors in the final version of L2 writing as a result of post-editing self-translated texts facilitated by comparison with a machine-translated version.

As for genre-specific effects, Chung and Ahn (2022) discovered that narrative writing had significantly higher improved accuracy scores than argumentative writing. However, MT assistance improved accuracy measures for both genres. Concerning moderator effects of

proficiency level, results indicate that language learners with a higher L2 level benefited more from MT assistance regarding accuracy than those with a lower proficiency level (Chung & Ahn).

Overall, while some studies report clear benefits of MT in enhancing accuracy, others yield mixed results or demonstrate specific contexts where MT aids in error reduction when writing in a second language.

2.2.3.4 Fluency

The impact of MT use on text length (often referred to as ‘fluency’) in L2 writing, too, varies across studies. Kol et al. (2018) observed that students wrote significantly more words when using Google Translate compared to unaided writing, with longer sentences and words. Tsai (2019) noted a similar trend, with L1 automatically translated texts containing more words compared to self-written L2 texts. Conversely, Chung and Ahn (2022) did not ascertain any significant changes in fluency. These findings indicate that MT use may result in an increase in text length in L2 writing, although this is not always the case.

2.2.3.5 Perceptions

Chung and Ahn (2022) found that over 90% of participants considered MT beneficial for L2 writing. Most reported using it for vocabulary (54%), sentence variety (19%), and grammatical accuracy (15%). Time efficiency, convenience (both 40%), and increased confidence (15%) were key reasons for use. Open-ended responses highlighted advantages like convenience, vocabulary support, grammar checking, and basic sentence creation, while disadvantages included grammatical errors, unintended meanings, awkward phrasing, and concerns about hindering language learning. Kol et al. (2018) similarly reported that 75% of students used MT primarily to look up words and phrases, with few translating entire sentences or texts. In Lee (2020), 88% of students reflected that MT positively impacted their writing. While initially sceptical, most were surprised by its improved accuracy and found it particularly useful for selecting precise vocabulary and authentic expressions. Around 70% indicated they would continue using MT for vocabulary support, with several noting that MT provided more contextually appropriate word choices than dictionaries. Finally, Tsai (2019) found general satisfaction among students using Google Translate for writing tasks. Learners believed it

helped them select more appropriate words and improved the overall quality and completeness of their writing.

2.2.4 Main findings of studies on the impact of MT use on L2 speaking

Research on the impact of MT use on L2 speaking has primarily focused on improvements in pronunciation. The perceived utility of MT tools as discussed in Rushton (2022) and Sakamoto (2022) can be found in Section 2.2.4.2.

2.2.4.1 Pronunciation

As mentioned in the theoretical preliminaries (Section 2.1), pronunciation is considered a key aspect of speaking skills. Therefore, it is a frequently researched feature in studies measuring the impact of MT on L2 speaking. He and Cardoso (2021) and Khademi and Cardoso (2022), for instance, investigated the effect of the use of TTS and ASR functionalities provided by online translators, such as Microsoft Translate and Google Translate, on the pronunciation of English past tense *-ed* endings by EFL students.

In terms of phonological awareness, the results of both studies indicated that the use of the online translators' speech technologies improved the participants' awareness of the *-ed* allomorph. Additionally, one study showed that the treatment improved the overall level of phonological awareness (He & Cardoso, 2021). Khademi and Cardoso (2022) noted, however, that the improvements were limited to the awareness of the pronunciation of /id/ and /t/, but not /d/.

The second stage of pronunciation acquisition according to Celce-Murcia et al. (2010), namely phonemic discrimination, in this instance of the three allomorphs (/t/, /id/, and /d/), was also improved by the treatment in both studies. Khademi and Cardoso's study (2022) also reported improvements in the discrimination of past and non-past verb forms. However, this change was not observed in He and Cardoso's study (2021), despite the similar research design.

Improvements in terms of spoken output were especially observed in He and Cardoso's study (2021). Participants' pronunciation of past *-ed* endings significantly improved in their study. Surprisingly, the treatment did not have any effect in the analogous study by Khademi and Cardoso (2022), except on the production of the /id/ allomorph.

Overall, these findings suggest that the pedagogical use of TTS and ASR technologies can benefit EFL students' pronunciation, particularly in regard to the English past *-ed* allomorph. The use of these MT features has been shown to increase their awareness of this allomorph and improve their ability to distinguish between the three pronunciation options (/t/, /id/, and /d/). While He and Cardoso (2021) observed significant improvements in the participants' oral production, Khademi and Cardoso (2022) reported limited effects, mainly on the production of the /id/ allomorph.

2.2.4.2 Perceptions

Although the studies by Rushton (2020) and Sakamoto (2022) did not directly investigate the impact of MT use on L2 speaking performance, they do provide valuable insight into learners' perceptions of these tools during speaking-related activities. Both studies reported generally positive attitudes towards the use of MT. Rushton found that learners believed DeepL made their English sound more engaging, helped them acquire new vocabulary and expressions, and increased their motivation to participate in oral communication classes. Moreover, most participants viewed the tool as a legitimate learning aid rather than a form of academic dishonesty. Similarly, participants in Sakamoto's study described DeepL as easy to use, useful, and that the tool improved their English in some way.

2.3 Impact of generative AI use on L2 production

Compared to the growing body of research on machine translation tools like DeepL and Google Translate, the academic exploration of generative AI applications in L2 learning is still in its early stages. Therefore, this section takes a thematic rather than systematic approach, by summarising trends and preliminary findings from recent studies.

A clear pattern in the existing literature is that almost all research in this field has been conducted in EFL contexts (Ali et al., 2023; Zhang & Huang, 2024; Huang & Mizumoto, 2024; Yamaoka, 2024; Tang, 2025; Laksana et al., 2024). Moreover, many of these studies were situated in university settings (Yamaoka, 2024; Tang, 2025; Laksana, 2024).

Most studies to date have relied on students' perceptions of AI tools like ChatGPT, typically gathered through questionnaires, surveys, or interviews (Ali et al., 2023; Huang & Mizumoto,

2024; Yamaoka, 2024; Tang, 2025). A few studies have also incorporated teachers' viewpoints, offering a more holistic perspective on AI's role in the classroom (e.g. Ali et al., 2023).

A recurring theme in these studies is learner motivation. Ali et al. (2023) reported that students generally held positive attitudes towards using ChatGPT for reading and writing tasks, while reactions to its use for listening and speaking were more neutral. Additionally, participants felt that ChatGPT supported their learning independence and boosted their self-confidence. Similarly, Huang and Mizumoto (2024) found that ChatGPT positively influenced students' Ideal L2 Self and L2 Learning Experience, though it did not affect their Ought-to L2 Self motivation. The tool also significantly enhanced students' writing self-efficacy, though the authors stressed the importance of structured guidance to mitigate risks such as plagiarism. Yamaoka (2024) likewise observed that ChatGPT reduced learners' anxiety, increased motivation, and was perceived as a valuable support tool in language learning. In line with these findings, Laksana (2024) concluded that generative AI tools positively impacted student motivation and engagement, fostering greater interest in learning English.

Regarding the impact on language learning outcomes, Zhang and Huang (2024) demonstrated that AI chatbots (though not ChatGPT specifically) could improve students' receptive and productive vocabulary knowledge. Other studies, while not experimental, gathered student perceptions suggesting that AI tools can enhance grammar and vocabulary acquisition (Ali et al., 2023) and improve writing quality and efficiency in areas such as language optimisation, content generation, and strategic support (Tang, 2025). However, these perceptions remain largely untested in controlled experimental studies, meaning concrete evidence of language proficiency gains is still limited.

Although most research has focused on AI's role in L2 writing, Laksana et al. (2024) addressed its perceived influence on speaking performance. Students in this study reported improvements in fluency and vocabulary while using generative AI tools, though, again, these findings were based on subjective impressions rather than empirical measurements. One specific advantage identified by participants was the provision of personalised feedback and error correction (Laksana et al., 2024). However, the potential risk of overreliance on AI tools was also raised, underlining the importance of carefully integrating these technologies into language education.

Taken together, these early studies underline both the promise and the challenges of generative AI for L2 learning. While learners recognise AI's motivational and supportive benefits, further empirical research is needed to substantiate claims about its effects on language proficiency and performance.

2.4 Research questions

Despite the growing interest in digital tools for language learning, several notable gaps remain in the current literature. First, while numerous studies have investigated the use of machine translation tools such as DeepL and Google Translate for both L2 writing and, to a lesser extent, speaking, little attention has been paid to how these tools compare with newer generative AI applications like ChatGPT in the context of L2 speaking performance. To date, most research on generative AI in language education has focused on learners' and teachers' perceptions or attitudes (Ali et al., 2023; Huang & Mizumoto, 2024; Yamaoka, 2024; Tang, 2025), rather than on objectively measuring the impact of these tools on spoken language outcomes.

To the best of our knowledge, no empirical studies have examined how the use of generative AI tools influences L2 speaking performance, particularly in terms of product-oriented outcomes such as fluency and lexical complexity. Existing research tends to rely on self-reported improvements or learner satisfaction. Therefore, there is a gap in evidence-based assessments of actual L2 performance following AI-assisted preparation.

The present study aims to address these gaps by systematically comparing three different types of language tools – traditional (online) dictionaries, machine translation, and generative AI – for L2 speaking preparation. In doing so, it seeks to provide insights into how these tools affect learners' oral production.

Specifically, this study investigates (1) how the use of these tools influences L2 learners' speaking performance, as measured through fluency, lexical complexity, and lecturer evaluation, and (2) how learners perceive and experience using these tools when preparing for an oral presentation in English. In addition (3), it seeks to discover how language students generally use these tools, and what their perceptions of their use are. By combining quantitative analysis of speaking performance with survey data on learner experiences, this research

contributes to a more comprehensive understanding of how digital tools can support L2 speaking development.

The following research questions guide the study:

- RQ1: To what extent does the use of dictionaries, a machine translation, or generative AI tool influence L2 learners' speaking performance in terms of fluency, lexical complexity, and lecturer evaluation?
- RQ2: How do learners experience and evaluate the use of these tools when preparing for an L2 speaking task?
- RQ3: How do learners generally use and experience these tools for language support in their L2?

Based on the available literature, the following hypotheses are proposed for each research question:

- RQ1: Drawing on studies from L2 writing contexts, it is hypothesised that machine translation tools facilitate the use of more sophisticated vocabulary, reflected in a lower proportion of high-frequency words and a higher proportion of academic and off-list words. These expectations are based on findings by Chung and Ahn (2022), Kol et al. (2018), and Tsai (2019), who reported improved lexical profiles in MT-assisted writing tasks. However, as no prior studies have experimentally examined the impact of MT or generative AI use on L2 speaking fluency or lecturer evaluations, no specific hypotheses are proposed for these variables.
- RQ2: Previous research into learner perceptions of MT use in writing contexts has reported largely positive experiences. Chung and Ahn (2022) found that over 90% of students considered MT beneficial, particularly for vocabulary lookup and time efficiency. Similarly, Kol et al. (2018), Lee (2020), and Tsai (2019) documented favourable learner attitudes toward MT tools for vocabulary support, contextual appropriateness, and improved writing quality. Therefore, it is expected that learners will report positive experiences using MT tools for L2 speaking preparation, especially regarding vocabulary support and convenience. Due to the lack of

experimental research into students' perceptions of generative AI tools in L2 speaking preparation, no hypothesis is formulated for this tool.

RQ3: Literature on MT use in L2 writing shows that learners predominantly use these tools for vocabulary and expression lookup (Chung & Ahn, 2022; Kol et al., 2018), grammar checking, and to improve text completeness and appropriateness (Lee, 2020; Tsai, 2019). These studies also highlight time efficiency and ease of use as important motivators. In line with these findings, it is expected that learners will report frequent use of MT tools for vocabulary and expression lookup in their general L2 use. Given the limited empirical research on generative AI tools and learner habits beyond writing contexts, no specific hypothesis is proposed for generative AI or dictionary use in this area.

Each research question is addressed in a separate chapter of this study, with the findings presented in Chapters 4, 5, and 6, respectively.

3 METHODOLOGY

In this experimental study, 30 L2 English students (3.1) gave group presentations on a controversial topic (3.2) using different language tools (3.3). These presentations were recorded and transcribed (3.4). Participants were also invited to complete a survey regarding their experience in preparing for the presentations and their overall experience with online language tools (3.5). Both the presentations and survey responses were analysed quantitatively and qualitatively (3.6).

3.1 Participants

The present study comprised 30 participants, with 27 females and 3 males. The participants were between the ages of 18 and 30, with an average age of 19.47 years and a median age of 19. All participants were Dutch speakers, although some also spoke additional languages at home, including Russian, Thai, Hindi, Spanish, and Arabic.

In accordance with the majority of prior studies examining the integration of language technology in education, the present study's participants were EFL university students (Chung & Ahn, 2022; Kol et al, 2018; Lee, 2020; Tsai, 2019; Rushton, 2022; Sakamoto, 2022; Yamaoka, 2024; Tang, 2025; Laksana, 2024). All participants were enrolled in the second year of the Bachelor of Arts in Applied Linguistics programme at Ghent University, specialising in English. Specifically, they were all taking the *Engels:Taalvaardigheid II* course, an English course focusing on writing and speaking skills. It should be noted that two participants were following a bridging programme. The participants were divided into three groups (hereinafter called 'language tool groups') based on the day they were following the course.

Exchange students enrolled in the course were excluded from the study due to differences in their first languages and academic backgrounds. Consequently, their recordings were excluded from the analysis and they were not invited to participate in the survey.

3.2 Speaking task

For the first part of the study, students were asked to participate in group presentations as part of a classroom activity. Assignment sheets detailing the requirements and guidelines for each group were provided to the students. These materials are available in the Appendix (A).

Each group of three to four students selected a controversial topic from a list of 52 issues provided by their lecturer. The following list details the issues that were selected by the students. A shorter name for each theme is provided between brackets to facilitate reference to the issues in subsequent parts of the study.

- Does personal tech encourage asocial behaviour, and if so, what can be done to address the problem? (Personal tech)
- Is taking a gap year a useful thing to do between finishing your secondary school studies and taking up a place at university? (Gap years)
- Zoos may be good for educational purposes but it's unnatural for animals to live in cages. Discuss. (Zoos)
- An individual who is elderly or terminally ill should have the right to choose when his/her life should end. To what extent do you agree with this statement? (Euthanasia)
- TikTok and similar platforms are having a serious impact on the attention spans of young people. (TikTok)
- Obesity is a debilitating condition, but sufferers are often blamed for a lack of self-control. This kind of reasoning is unacceptable. (Obesity)
- Beggars on the street should be ignored. More than likely any money they receive will end up in the pockets of criminal gangs. Discuss. (Beggars)
- Abortion remains an extremely divisive issue in society, but ultimately it is an issue of whether or not a woman has agency over her own body. Discuss. (Abortion)

The students started to prepare their presentations in class. They were allowed to continue preparing at home. The allotted time for this task was one week. The participants were invited to document the time they had spent preparing for this task.

Each student was allotted three to four minutes of speaking time during their group presentation. The presentations consisted of three primary sections. However, the students were not instructed on how to divide the different sections among the speakers. The introduction comprised a hook to capture the audience's attention, a concise topic introduction, and a clear

thesis statement. The body of the presentation then presented well-structured arguments that supported the thesis, while the conclusion summarised the key points and provided a strong closing statement.

To enhance their delivery, students were permitted to use different presentation aids. They were allowed to project slides, use the blackboard, or play audio or video materials.

It is important to note that these presentations had no impact on the students' course marks. Conversely, the presentations provided an effective opportunity for students to prepare for their oral examinations, in which they were also required to argue their position on a controversial issue.

3.3 Language tools

The students were divided into three groups based on the day they followed the course. Each group was allotted a different type of tool for language support during their preparation at home. It should be noted that all students were permitted to use other (online) resources in addition to their assigned language tool to gather information about their topic and formulate their arguments.

The first group, following the course on Monday, was instructed to only employ (online) dictionaries for language assistance. Participants in this group were permitted to use translation dictionaries, such as Van Dale, as well as explanatory dictionaries, such as Collins Dictionary, Oxford Dictionary, Cambridge Dictionary, and Merriam-Webster. This group will henceforth be designated the *DI-group*. It was emphasised that the use of additional language tools such as DeepL, Google Translate, Reverso Context, ChatGPT, or QuillBot was prohibited.

The students enrolled in the course on Tuesday were requested to only use a machine translation tool for language support during their preparations. They were all asked to use DeepL Translate. The decision to restrict the use of machine translation to a single tool was made to minimise variations in the quality of the output produced by different translation tools. The decision to use DeepL as opposed to Google Translate, another widely employed machine translation tool, was made because it allows the user to choose from multiple phrasing options. The *MT-group* (henceforth referred to as such) was thus instructed not to use Google Translate, online dictionaries, Reverso Context, and generative AI tools such as ChatGPT and QuillBot.

Finally, the Friday group engaged with a generative AI tool, namely ChatGPT. Once more, the decision was taken to instruct the students to use a single, designated tool, thus ensuring that the quality of the AI experience would be uniform for all participants. Participants in the *AI-group* were instructed not to employ online dictionaries, Reverso Context, QuillBot, Google Translate, or DeepL during their preparation.

Table 3 provides an overview of the three language tool groups, including the number of students and group presentations, in addition to the topics selected by each group for their presentations. The table contains the shortened names of the topics, the full versions of which can be found under 3.2.

	DI-group	MT-group	AI-group
Number of students	9	7	14
Number of group presentations within the language tool groups	3	3	4
Topics chosen by each group, along with the number of students in that group ³	1. Personal tech (3) 2. Gap years (3) 3. Zoos (3)	1. Zoos (3) 2. Euthanasia (1) 3. TikTok (2)	1. Obesity (4) 2. Beggars (3) 3. Zoos (4) 4. Abortion (3)

Table 3: Overview of the numbers of students in each language tool group, and the topics they selected for their group presentations

For the purposes of ease of reference and analysis, presentations will henceforth be labelled as follows: DI2 for the second presentation of the group who were asked to use dictionaries for language support, AI3 for the third presentation of the group that prepared their presentations with ChatGPT. Participants were anonymised as follows: MT1.3 is the third student of the first MT-group presentation, and AI4.2 is the second student of the fourth AI-group presentation.

³ As mentioned in section 3.1, exchange students participating in the course were excluded from the study. Nevertheless, they did participate in the presentations, as these were a classroom activity. This is the reason why some groups comprise a single participant (e.g. MT2) and some transcriptions contain gaps (e.g. the AI2 presentation, see Appendix E).

3.4 Recordings and transcriptions

During their presentations, participants were invited to wear a lavalier microphone. This initially caused some unease and fiddling among the participants, but was considered to be the least intrusive way of recording their speech.

Subsequently, the presentations were automatically transcribed using Whisper. Afterwards, they were manually checked. For the purpose of the current study, we focused on speed fluency (see 2.1.1), and therefore, pauses were not included in the transcriptions. Nevertheless, extensive disfluencies were included to ensure an accurate representation of the number of words and, consequently, the speech rate. For instance, when one of the participants (DI2.1) said the following: “[...] but it can also... it can also give the student a new interest”, this was not simply transcribed as “but it can also give the student a new interest”, as that would distort the number of words and therefore also the speech rate.

3.5 Survey

After their presentations, the participants were sent a link to an online survey. The objective of this Qualtrics survey was to ascertain two aspects regarding the participants' experiences with (online) language tools. Therefore, the survey was divided into two sections. The survey questions can be found in the Appendix. They are based on the surveys conducted by Chung and Ahn (2022) and Niño (2020).

The first section of the survey was designed to assess the participants' experience of using their designated tool (i.e. dictionaries, DeepL, or ChatGPT) during the preparation of their presentation assignments. Consequently, the questions in this section differed for each language tool group. In the second section of the survey, students were invited to share their overall experience with online dictionaries, machine translation tools, and generative AI tools. All students therefore filled in the same questions.

Finally, the study incorporated a series of demographic questions. The objective of this last section was to ascertain whether there were considerable differences in age, gender, or linguistic background.

3.6 Analysis

The analysis of both the presentation data and the survey results was conducted using Microsoft Excel. Descriptive statistics were calculated to summarise speaking performance in terms of fluency (presentation duration, word count, and speech rate) and lexical complexity. Lexical complexity was operationalised through measures of lexical density, lexical diversity (type-token ratio), and lexical sophistication (proportions of K1 and K2 words, off-list words, and words from the Academic Word List). These lexical measures were computed using LexTutor (www.lextutor.ca/vp/eng/), following the approach of earlier studies such as Tsai (2019) and Kol et al. (2018). Due to time constraints, no analyses were performed on grammatical accuracy, as all participants were intermediate L2 English speakers, or grammatical complexity.

To identify statistically significant differences between the three language tool groups, Kruskal-Wallis tests were carried out for each quantitative variable. Pivot tables in Excel facilitated the organisation and comparison of descriptive data, and the results were visualised using bar charts and box plots to clearly illustrate group differences.

In addition to the quantitative analysis, a limited qualitative analysis was performed on the open-ended survey responses. Both the prompts used by the AI-participants and the advantages and disadvantages of the assigned language tools, as described by the participants, were categorised into thematic groups. Because of limited time, the answers to the open-ended questions regarding participants' tool preference were not analysed.

4 SPEAKING PERFORMANCE

In this first results section, the participants' presentations will be analysed regarding fluency (4.1), in terms of their evaluation by the lecturer (4.2), and with regard to lexical complexity (4.3). In each subsection, the results will be compared across the three language tool groups. In 4.4, the impact of two presentation aids on the variables mentioned in sections 4.1 to 4.3 will be discussed.

4.1 Fluency

All participants were instructed to speak for a duration of three to four minutes during their group presentations. Participants were not encouraged to learn a script of a particular number of words by heart; however, the transcriptions allowed for the calculation of the word count of each participant's speaking part. Consequently, the speech rate could also be calculated and subsequently compared across participants and language tool groups.

The mean duration of speech by each participant was 3 minutes and 45 seconds, with an average word count of 469.8. This equates to a speech rate of 124.25 words per minute. This rate is notably lower than the average speaking speed in British English, which is reported to be 198 words per minute (Wang, 2021).

Table 4 presents the results for each language tool group. The findings indicate that participants who had utilised machine translation for their presentations exhibited the highest word count, duration, and speech rate. In contrast, the DI-group demonstrated the lowest word count, the shortest duration, and the slowest average speech rate.

Group	Word count	Duration	Speech rate
DI	413.44	00:03:32	117.26
MT	562.43	00:04:04	133.92
AI	459.71	00:03:43	124.81

Table 4: Overview of the average word count, duration, and speech rate (words per minute) for each language tool group

In order to obtain a more comprehensive understanding of the dispersion of the effects of the different language tools, the speech rate results per group are provided in a box plot below (Figure 2). However, a Kruskal-Wallis test showed that the differences in word count,

duration, and speech rate across the three groups were not significant ($p = 0.250$, $p = 0.671$, and $p = 0.724$, respectively).

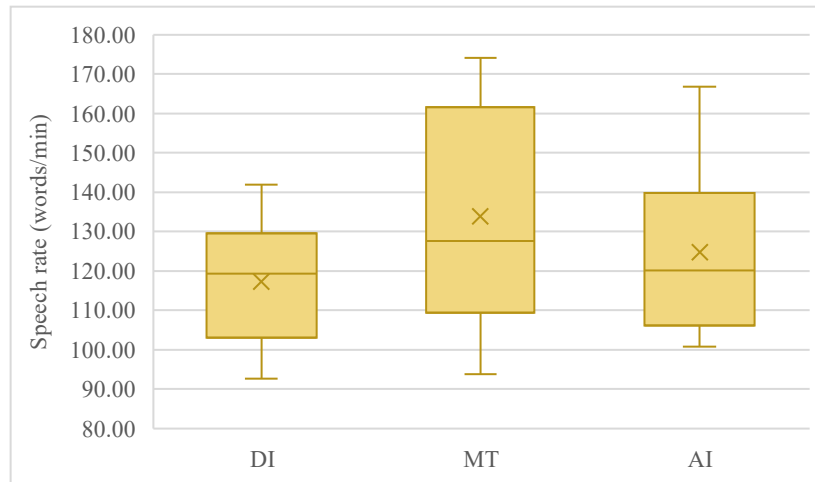


Figure 2: Speech rate (words/min) dispersion per language tool group

4.2 Lecturer evaluation

Given that these group presentations were not only the subject of this research paper but also primarily a classroom activity, the students were provided with oral feedback immediately after their presentations. In addition, the lecturer provided a performance evaluation on a scale of 1 to 20. It is important to note that these marks were not disclosed to the students, nor did they have an impact on the students' final mark for the course – the marks were exclusively shared with the researcher. It is noteworthy that the lecturer had known the students for a minimum of eight weeks before the presentations. Consequently, the marks may be perceived as slightly biased, though the lecturer has extensive experience in grading students' speaking skills.

The mean mark awarded was 14.33 out of 20. The MT-group demonstrated the highest average score of 15.29. The students who had prepared their presentations with ChatGPT received an average score of 14.71. The DI-group demonstrated the poorest performance according to the lecturer, achieving an average mark of 13.00 out of 20. The dispersion of these results is illustrated in Figure 3. A Kruskal-Wallis test revealed that these results are statistically significant ($p = 0.020$). Pairwise comparisons of these results indicated that the differences in the evaluation are significant between the DI- and AI-group ($p = 0.005$), and between the DI- and MT-group ($p = 0.009$).

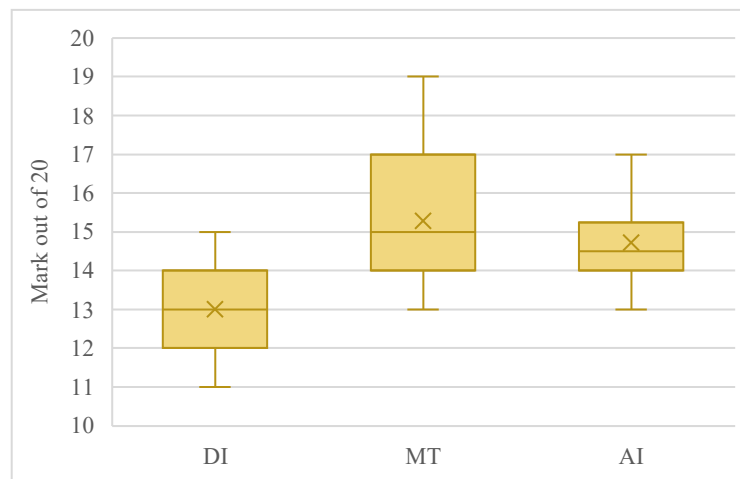


Figure 3: Mark dispersion per language tool group

4.3 Lexical complexity

The transcriptions of the presentations were subsequently processed through the VocabProfiler by Lextutor⁴. This online tool enables comparison of texts or, in the case of the current study, transcriptions, in terms of lexical complexity (i.e. lexical density, diversity, and sophistication). The site provides the user with the lexical diversity (type-token ratio) and lexical density (content words/total) of each transcription. The results for these two variables per language tool group are shown in Figure 4. While the average type-token ratio for the MT and DI groups is equivalent (0.41), that for the AI group is slightly higher at 0.45. The Kruskal-Wallis test suggested no statistically significant difference between the groups ($p = 0.179$). However, the difference in lexical diversity was shown to be significant between the MT- and AI-group ($p = 0.04$). Furthermore, the AI-group demonstrated the highest lexical density score (0.49). The average lexical density of the MT-group was 0.48, and the DI-group presentations were the least lexically dense with a score of 0.46. However, the results regarding lexical density were not statistically significant ($p = 0.336$).

⁴www.lextutor.ca/vp/eng/

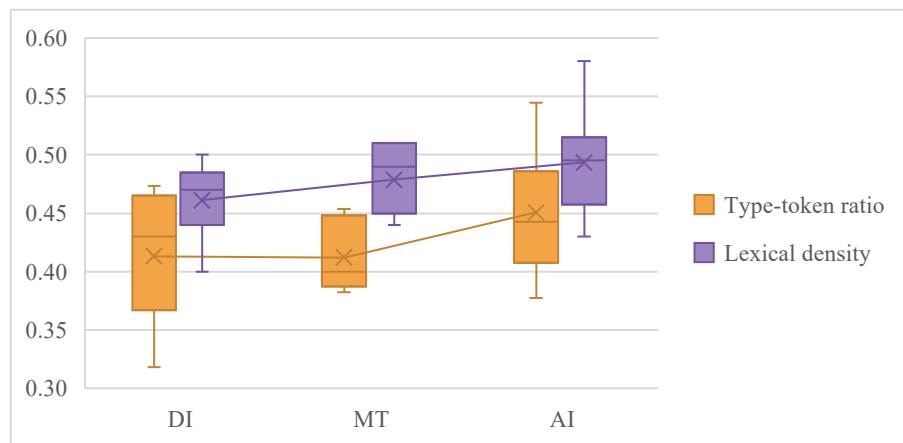


Figure 4: Type-token ratio and lexical density dispersion across the three language tool groups

A third component of lexical complexity according to Housen et al. (2012) is lexical sophistication. In this study, this was also quantified using Lextutor's VocabProfiler. It provides insights into the proportion of words from the K1, K2, AWL and off-list categories contained within each transcription. In other words, the tool provides insight into the proportion of words belonging to four distinct frequency levels: the list of the 1000 most frequent word families (K1), the second 1000 most frequent word families (K2), the Academic Word List (AWL), and words that do not appear in the other lists (off-list). The results demonstrate that the presentations by the DI-group exhibit the highest average proportion of K1 words. In contrast, the MT-group and AI-group presentations contain a lower proportion of K1 words, which is counterbalanced with higher proportions of the other categories. No significant differences across all three groups were found according to the Kruskal-Wallis test ($p = 0.088$). Nonetheless, pairwise comparisons reveal that the differences in the proportion of K1 words are statistically significant between the DI- and AI-group ($p = 0.017$), and between the DI- and MT-group ($p = 0.046$). Additionally, the findings indicate that the MT-group demonstrates the highest proportion of off-list words (not significant across the three groups, $p = 0.069$; significant difference with the DI-group – $p = 0.008$), while the AI-group exhibits the highest percentage of academic vocabulary (not significant across the three groups, $p = 0.090$; significant difference with the DI-group – $p = 0.013$). Figure 5 provides a visualisation of the mean lexical profiles.

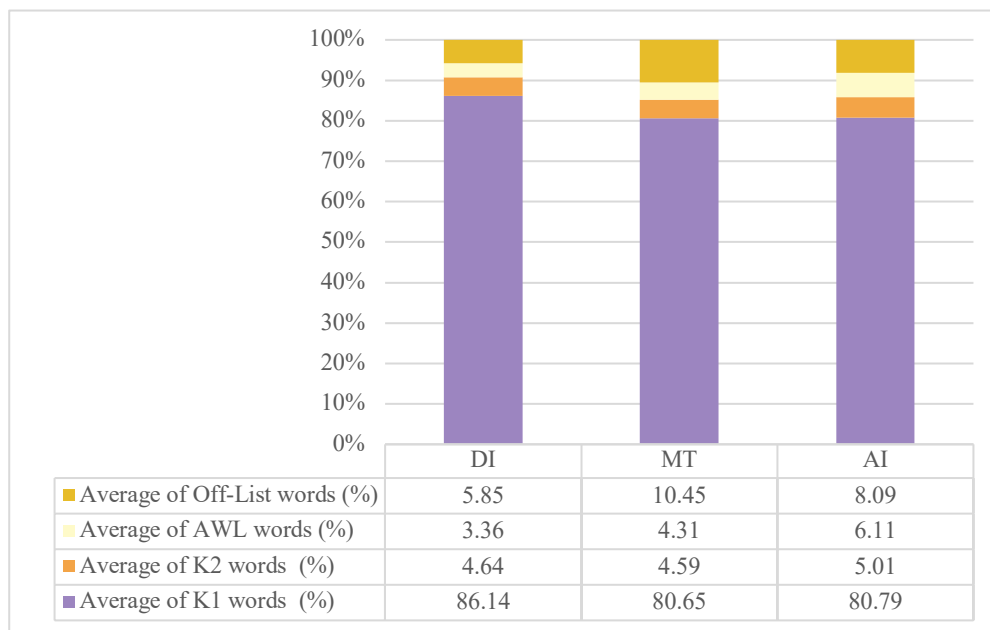


Figure 5: Average vocabulary profile per language tool group

4.4 Presentation aid use

As mentioned in section 3.2, participants were allowed to use different presentation aids. Participants could opt for any of the following options: projecting slides, utilising the blackboard, or playing audio or video materials. However, the only aid that was chosen were slides.

In total, 21 out of the 30 participants (or 6 out of the 10 groups) made use of slides during their presentations. Table 5 shows the average results of the parameters discussed in sections 4.1 to 4.3, depending on the use of slides. Participants who employed slides spoke for an average of 7 seconds longer. Furthermore, their usage of academic vocabulary was marginally lower, whilst their usage of off-list words was somewhat higher. However, a Mann Whitney U-test identified no significant disparities among these variables when comparing participants who employed slides with those who did not (see last row of Table 5).

Use of slides	Duration	Word count	Speech rate	Mark	Type-token ratio	Lexical density	K1 words (%)	K2 words (%)	AWL words (%)	Off-List words (%)
No	00:03:40	474.11	126.25	14.67	0.44	0.48	82.04	4.96	5.59	7.40
Yes	00:03:47	467.95	124.00	14.19	0.43	0.48	82.51	4.73	4.55	8.21
Total	00:03:45	469.80	124.67	14.33	0.43	0.48	82.37	4.80	4.86	7.97
<i>p-value</i>	0.904	0.904	0.973	0.657	0.940	0.973	0.994	0.656	0.597	0.551

Table 5: Average presentation fluency, marks, and lexical complexity depending on the use of slides

During the course of the presentations, the researcher observed that some participants were consulting notes. The subjects had not received any instructions as to whether this was permitted, but the question arises as to whether this has an effect on their delivery. Table 6 shows the average results regarding fluency, evaluation, and lexical complexity, contingent on the use of notes. The findings indicate that participants who presented without the use of notes spoke for an average of 19 seconds longer, and at an average speech rate of 14 words per minute faster than the participants who did use notes. In contrast, participants using notes during their presentation demonstrated an elevated type-token ratio (0.46 vs. 0.41, $p = 0.025$) and lexical density (0.49 vs. 0.47). The vocabulary profiles (i.e. the proportion of K1, K2, AWL, and off-list words) also present slightly more pronounced disparities than participants who did and did not use slides during their presentations. However, these differences were not statistically significant.

Use of notes	Duration	Word count	Speech rate	Mark	Type-token ratio	Lexical density	K1 words (%)	K2 words (%)	AWL words (%)	Off-List words (%)
No	00:03:53	502.47	130.71	14.47	0.41	0.47	83.20	4.65	4.08	8.07
Yes	00:03:34	427.08	116.77	14.15	0.46	0.49	81.28	5.00	5.88	7.83
Total	00:03:45	469.80	124.67	14.33	0.43	0.48	82.37	4.80	4.86	7.97
<i>p-value</i>	0.791	0.255	0.134	0.934	0.025	0.655	0.678	0.386	0.107	0.847

Table 6: Average presentation fluency, marks, and lexical complexity depending on the use of notes

5 PREPARATION PROCESS AND REFLECTIONS ON TOOL USAGE

After their presentations, the participants engaged in an online survey to reflect on their experience. In the first half of the survey, participants were asked to reflect on their experience of using their assigned tool for this particular presentation. In other words, each language group was presented with a different series of questions, contingent on the tool they had used to prepare their presentations. Initially, respondents were asked to specify the amount of time allocated to the preparation stage. In the subsequent series of questions, respondents were invited to evaluate their experience of preparing for their presentations using their assigned tool (i.e. dictionaries, DeepL, or ChatGPT).

5.1 Preparation time

The participants were asked to approximately measure, in minutes, the amount of time they allocated to the preparation of their presentations at home. The mean duration allocated for this activity was 116.67 minutes. In contrast, students who were instructed to use dictionaries required an average of 85 minutes. The MT-group allocated an average of 90 minutes to the task. The AI-group allocated the most time to this activity, with an average of 150.36 minutes. The differences in preparation time were not shown to be significant. The box plot (Figure 6) below shows the dispersion of the results per language tool group.

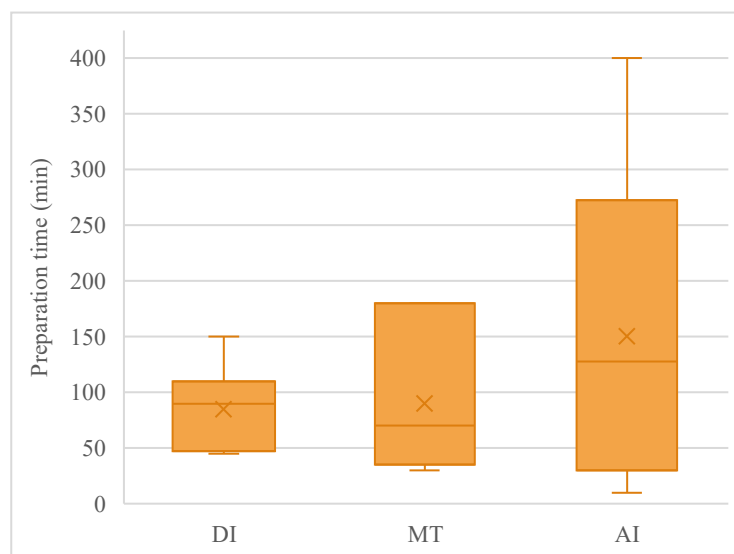


Figure 6: Preparation time (min) dispersion per language tool group

5.2 Language tool use

After indicating how much time they had spent preparing their presentations, participants were asked how they had used the (type of) tool they had been allocated. Naturally, questions differed slightly for each language tool group.

5.2.1 DI-group

Participants in the DI-group were asked which dictionaries they had used. Translation dictionaries were shown to be the most used – six out of the nine participants used Van Dale and three used mijnwoordenboek.nl. Many participants also used explanatory dictionaries – the Cambridge Dictionary was used most, chosen by four out of nine participants, the Oxford Dictionary by two and Merriam Webster by one participant. However, most participants used more than one dictionary. Four of them used a combination of translation and explanatory dictionaries. Three used only translation dictionaries and two used exclusively English explanatory dictionaries. Although the participants had been explicitly instructed not to use DeepL, one participant indicated that they had used this 'dictionary' to prepare their part of the presentation. This was not an option in the multiple-choice question, but the student had written it down in the 'Other:' option. One participant also wrote down that they had used Reverso Context, an online bilingual dictionary tool that provides translations of words and phrases in context. Both cases suggest that the difference between dictionaries and machine translation tools may not be fully understood by all participants. Figure 7 visualises these results.

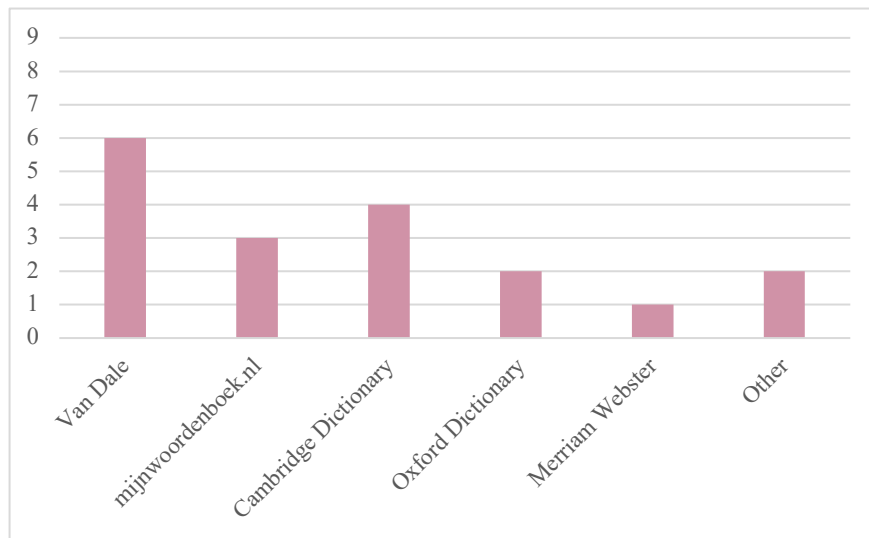


Figure 7: Overview of the dictionaries participants in the DI-group reported having used to prepare their presentations (total number of participants: 9)

The DI-group was also asked to specify the purposes for which they had used the dictionaries. The results of the study revealed that seven out of the nine participants in this group used dictionaries to look up words. Furthermore, only three participants employed dictionaries for looking up phrases, one participant employed them for expressions, and three students used them to look up pronunciation. Two participants selected the 'Other' option, with one participant using dictionaries to verify the spelling of words and the other to identify synonyms or alternative “ways to say things”. In Figure 8, these findings are summarised.

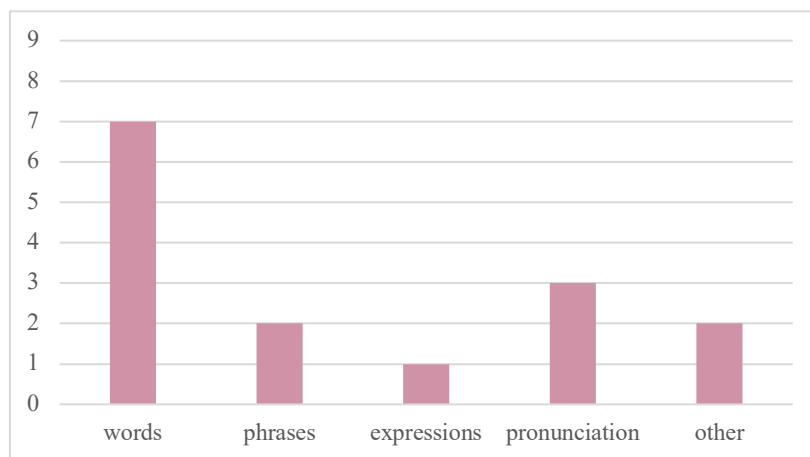


Figure 8: Overview of the types of information participants in the DI-group reported looking up in dictionaries for their presentations (total number of participants: 9)

5.2.2 MT-group

The MT-group was first asked to specify the type of input they had entered most frequently into the DeepL translation tool. The majority of participants (six out of seven) reported utilising the tool for the purpose of looking up words. Two participants reported using it to look up phrases, one participant to translate clauses, and two participants to translate sentences into English. No participants reported using DeepL to translate paragraphs for their presentations. See Figure 9 for an overview.

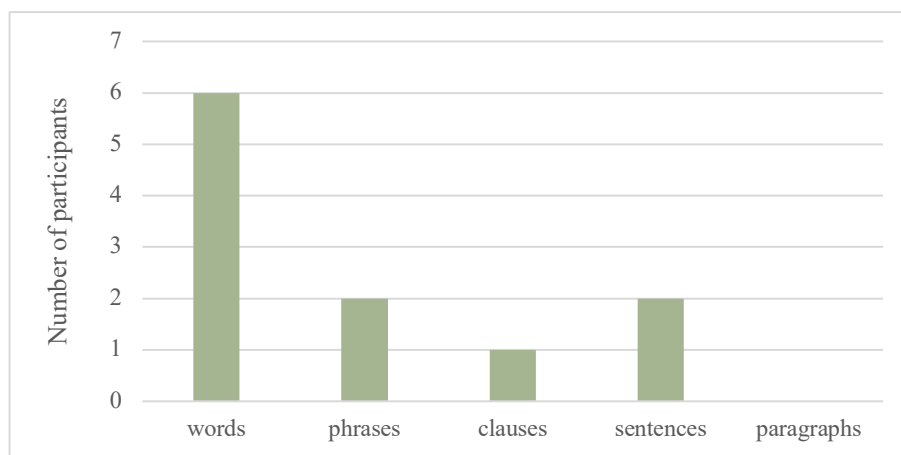


Figure 9: Type of input entered into DeepL by the MT-group (total number of participants: 7)

Subsequently, the participants in this language tool group were asked in which ways DeepL was beneficial for their preparation. The majority of the participants (six out of seven) used the machine translation tool to look up vocabulary. Three participants indicated that they had used DeepL to verify pronunciation during their preparation. One participant used it to verify spelling, and one participant to ensure grammatical correctness. Had the participants been required to produce written texts, these numbers would probably have been higher. No participants reported using DeepL to write various types of sentences. One participant reported using DeepL to ‘see different options for the same translation’, a unique feature of this machine translation tool. See Figure 10 for an overview.

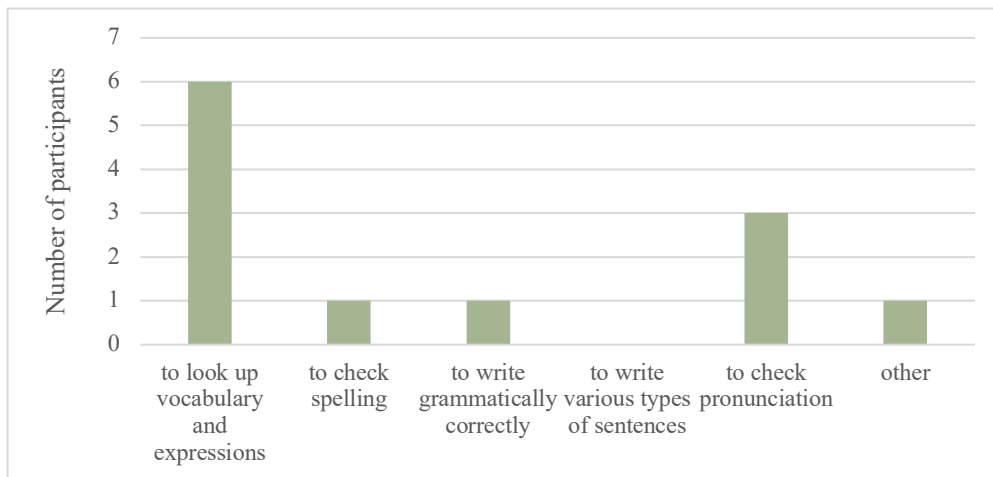


Figure 10: Ways in which DeepL was beneficial during the preparation of the presentation according to the MT-group (total number of participants: 7)

The MT-group was then invited to comment on the stage of their preparation at which they found the machine translation tool to be especially beneficial. The majority of participants (four out of seven) reported that DeepL was most beneficial during the composition of their respective parts of the group presentation. Two participants reported it to be especially helpful when practicing their part of the presentation. Finally, one participant reported that the tool was most valuable for brainstorming ideas.

The fourth question for this language tool group pertained to the extent to which the participants accepted DeepL's output. Most of the participants (four out of seven) reported that they had accepted the machine translation output 61 to 80 % of the time. See Figure 11 for an overview.

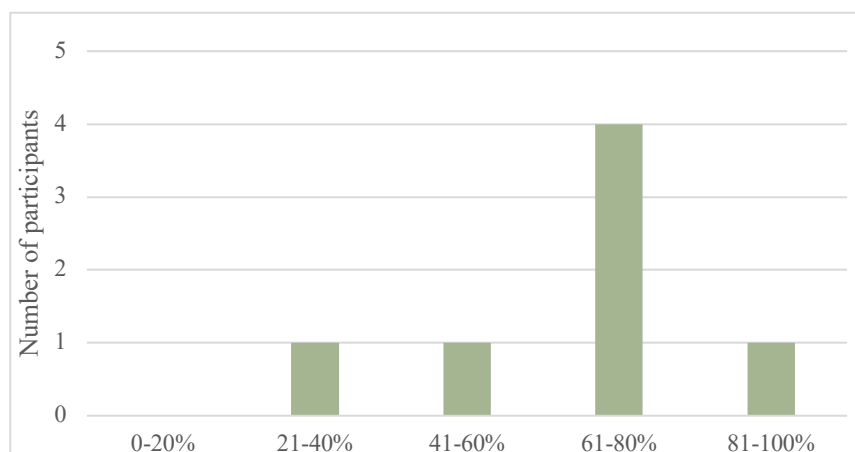


Figure 11: Acceptance rate of the DeepL output by the MT-group (total number of participants: 7)

When the participants were asked why they did not accept the output generated by DeepL when they did not, the following results were obtained. All participants reported that they did not accept the output when its contextual meaning was different from what they initially had envisaged. Additionally, one participant reported rejecting the DeepL output because it was grammatically incorrect, while another indicated being uncertain about the sentences.

4.2.3 AI-group

The students who had prepared their presentations using ChatGPT were firstly invited to enter the prompts they had used into a text slot in the survey. These prompts are located in the Appendix (C). The following three broad categories can be distinguished: linguistic support, content development, and presentation support prompts. The following table provides a selection of examples for each of the categories of prompts (see Table 7).

Linguistic support	Content development	Presentation support
How do you translate 'kant-en-klaar'? (AI1.4)	Can you give me some examples of wildlife sanctuaries around the world? (AI3.2)	Make a presentation about 20 minutes long about why or why not should we give money to beggars on the street (AI2.1, AI2.2, AI2.4)
Can you say 'what's more is that ...'? Or does it sound clunky? (AI1.1)	Could you give me some examples of physical consequences of being denied an abortion? (AI4.1)	
Is grocery store British English? (AI1.2)		

Table 7: Examples of prompts by participants in the AI-group divided into three categories

After entering their prompts into the survey, the participants were invited to articulate the ways in which they had found ChatGPT beneficial during the preparation stage. The vast majority of participants (13 out of 14) confirmed that the tool helped them gather ideas for their argumentation. About two-thirds of the participants (10) found it helpful to answer their questions. Seven participants found ChatGPT to be beneficial to generate a draft for their part of the presentation. One participant reported using it to give feedback on their work. These results are visualised in Figure 12.

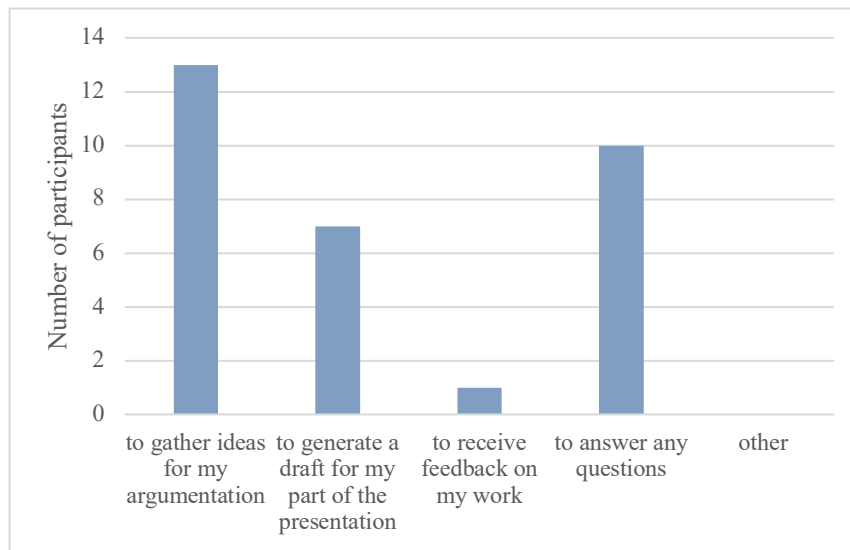


Figure 12: Ways in which ChatGPT was beneficial during the preparation of the presentation according to the AI-group (total number of participants: 14)

The next question posed to the AI-group aimed to ascertain the stage of the preparation process at which participants found ChatGPT to be particularly valuable. Almost two-thirds of the respondents (9) selected the option ‘When gathering ideas’. Five participants found the tool to be the most beneficial when writing their respective parts of the group presentation. Furthermore, no participants considered the tool to be the most helpful when practicing for their presentation. Figure 13 visualises these results in the form of a bar chart.

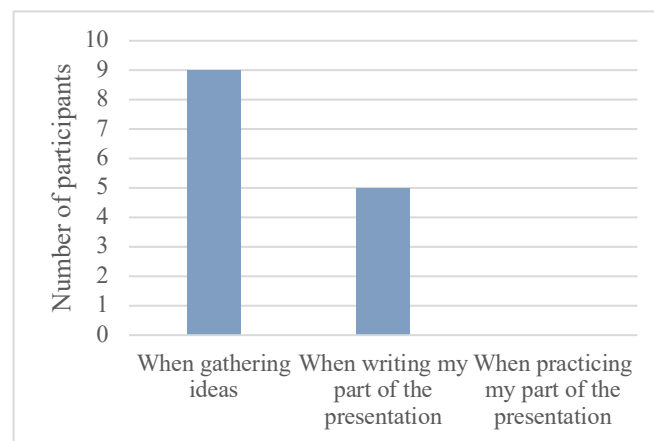


Figure 13: Stage at which participants in the AI-group found ChatGPT to be especially helpful (total number of participants: 14)

Students were then invited to indicate the frequency with which they accepted the output generated by ChatGPT. Approximately half of the participants (8 out of 14) indicated that they

accepted it 61 to 80% of the time. Six students reported that they were more inclined to accept the output (81 to 100 % of the time). Figure 14 provides an overview of the answers.

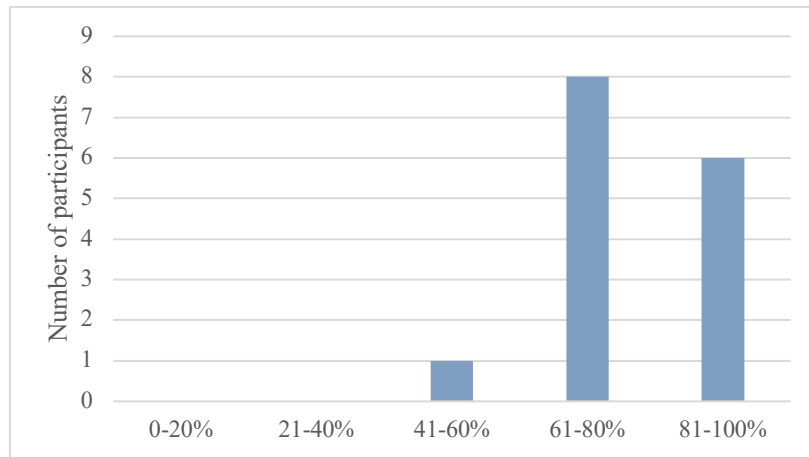


Figure 14: Acceptance rate of the ChatGPT output by the AI-group (total number of participants: 14)

The participants were also invited to provide a rationale for their decision to reject the ChatGPT output. The majority of participants (9 out of 14) did not accept the ChatGPT suggestions, as its contextual meaning differed from their initial expectations. Other prevalent rationales encompassed uncertainty regarding the output (five respondents), the perception that the tone was overly formal (four), and apprehension about their credibility if they would utilise the output (five). One participant responded that the information provided was not pertinent to their part of the presentation. See Figure 15 for the other answers.

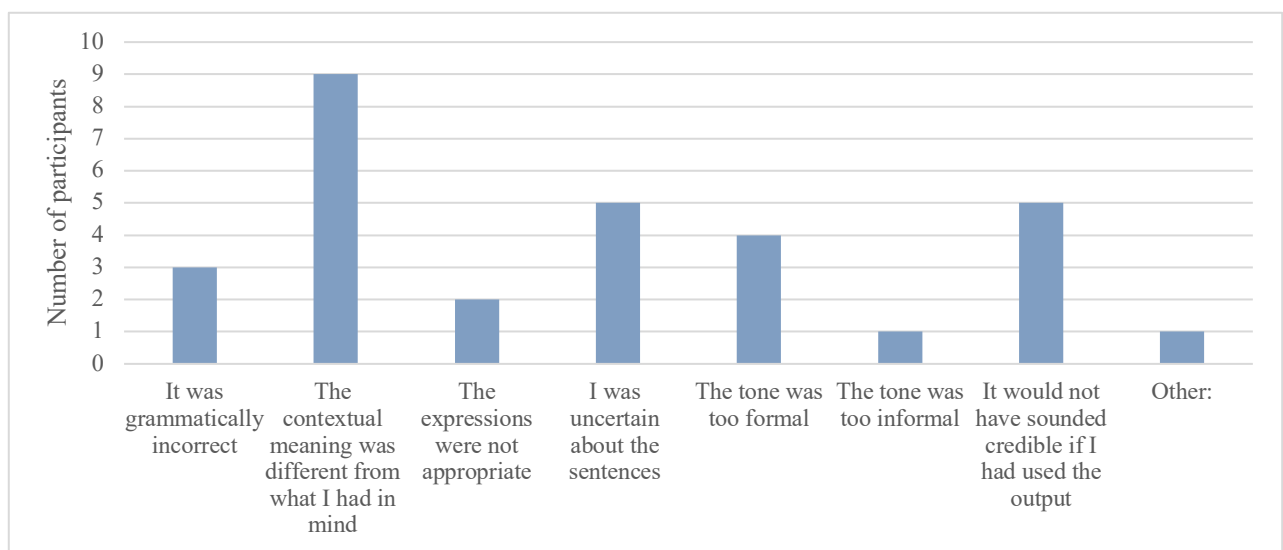


Figure 15: Reasons why the AI-group did not accept the ChatGPT output (total number of participants: 14)

Comparison

In this section, a brief comparison of the answers to two questions posed to both the MT-group and AI-group will be presented.

The first question posed to both groups pertained to the stage at which they found their assigned tool to be especially beneficial. As illustrated in Figure 16, the majority of participants in the AI-group (64.29%) found their tool, namely ChatGPT, to be most valuable for the generation of ideas for their presentations. Conversely, the MT-group placed the greatest value on DeepL during the writing stage of the preparation process, with 57.14% of respondents indicating this option. It is also noteworthy that only participants from the MT-group reported to use their tool when practising their part of the presentation. According to a Mann-Whitney U test, these differences are statistically significant ($p = 0.025$).

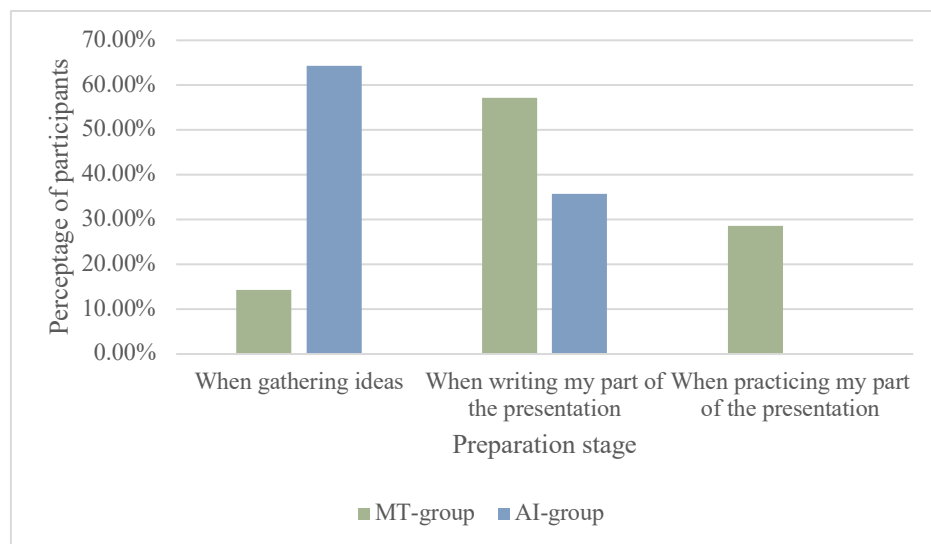


Figure 16: Stage at which participants in the MT-group and AI-group found their assigned tool to be especially beneficial

The second question that was posed to both groups aimed to gauge their acceptance rate of the tools assigned to them. As demonstrated in Figure 17, the MT-group and AI-group exhibited a high degree of acceptance of the output from their designated tool, with 52.14 % of both groups accepting the output 61 to 80 % of the time. In the AI-group, there was a higher proportion of students who demonstrated a greater propensity to accept the output of their tool (42.86 % vs. 14.29 % in the MT-group). In contrast, a small number of participants in the MT-group were more critical towards the DeepL output, by only accepting it 21 to 60 % of the time. However, these differences in acceptance rates were not shown to be significant ($p = 0.224$).

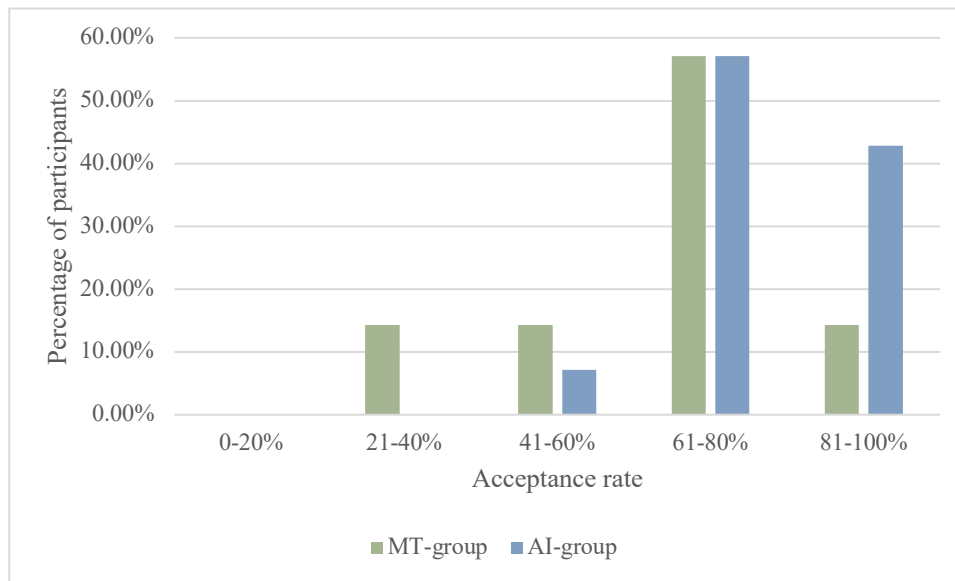


Figure 17: Acceptance rates of the DeepL and ChatGPT output by the MT-group and AI-group, respectively

5.3 Attitudes towards assigned tool

At the end of this part of the survey, where students were invited to provide feedback on their experience of preparing for the presentations, all language tool groups were asked how satisfied they were with their assigned tool and how useful they found it to prepare their presentations. This section will compare the responses provided by the three groups.

5.3.1 Perceived enjoyment

Initially, each group was requested to indicate how much they enjoyed using their assigned tool, i.e. dictionaries, DeepL, or ChatGPT, for the preparation of their presentations. The evaluation of these tools was conducted on a scale of one to five stars (see Appendix B). As demonstrated in Figure 18, both the DI-group and the AI-group expressed a high level of enjoyment with their assigned tool, with scores ranging between three and five stars, and an average of four stars in both groups. The mean average score achieved by participants who had used DeepL for the preparation of their presentations was 3.57. This indicates that, on average, these participants found this tool less enjoyable. However, according to a Kruskal-Wallis test, the differences between language tool groups are not significant ($p = 0.599$).

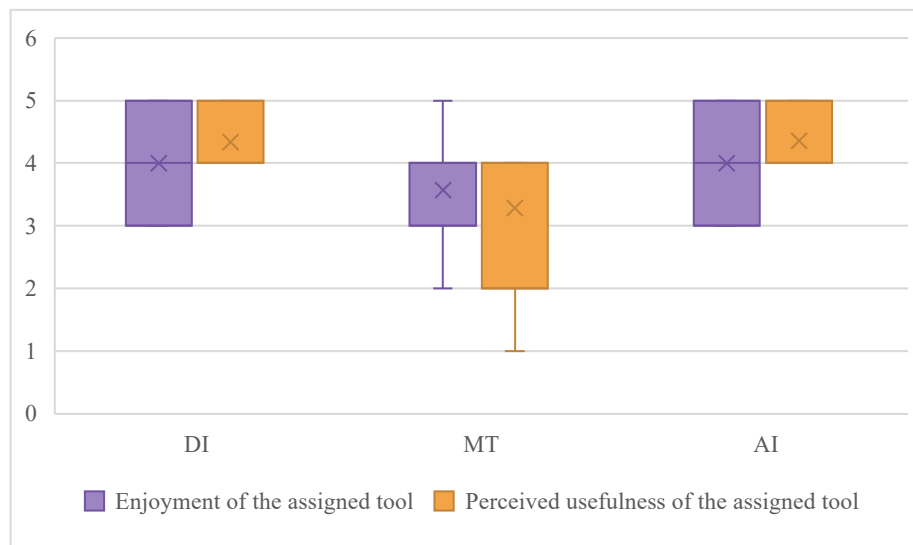


Figure 18: Perceived enjoyment and usefulness dispersion across the three language tool groups

5.3.2 Perceived usefulness

A second question aimed to assess the participants' attitudes towards their assigned tool by enquiring about its perceived usefulness. These results are also illustrated in Figure 18. For this question, respondents were once again invited to provide their answers on a scale of one to five stars. However, on this occasion, they were asked to rate how useful they found their assigned language tool to be for preparing this assignment. As with the previous variable, the mean ratings from the DI-group and AI-group were comparable, with average scores of 4.33 and 4.36, respectively. It is noteworthy that across both groups, all participants assigned a rating of four or five stars to this question. In contrast, students in the MT-group exhibited a comparatively diminished level of enthusiasm for their language tool – their scores ranged from one to four stars with respect to the perceived usefulness of DeepL, with a mean score of 3.29. These differences proved to be statistically significant ($p = 0.039$).

6 GENERAL EXPERIENCE WITH LANGUAGE TOOLS

The second half of the survey was designed to gauge the participants' general experience with dictionaries, machine translation, and generative AI tools. These questions were posed to all participants. This chapter will address how the students in this study usually employ the three aforementioned types of tools (6.1) and what their attitudes towards their use are (6.2).

6.1 Language tool use

6.1.1 Dictionaries

Participants were first asked about their experience with dictionaries. The first question in this section was the following: “Which dictionary (or dictionaries) have you used before?” The results of this survey are illustrated in the bar chart below (Figure 19). It was reported by all participants that they had previously used Van Dale. Furthermore, 28 out of the 30 participants had also worked with mijnwoordenboek.nl. The Cambridge Dictionary and the Oxford Dictionary were the most frequently cited explanatory English dictionaries, with 16 and 18 participants having used these, respectively. Six participants reported having worked with 'Other' dictionaries before. These included the Longman Dictionary (mentioned four times), the Thesaurus (once), Reverso Context (once), and DeepL (once). These last two answers once more indicate that not all participants had a clear understanding of the definition of a dictionary, as Reverso Context and DeepL are not purely dictionaries.

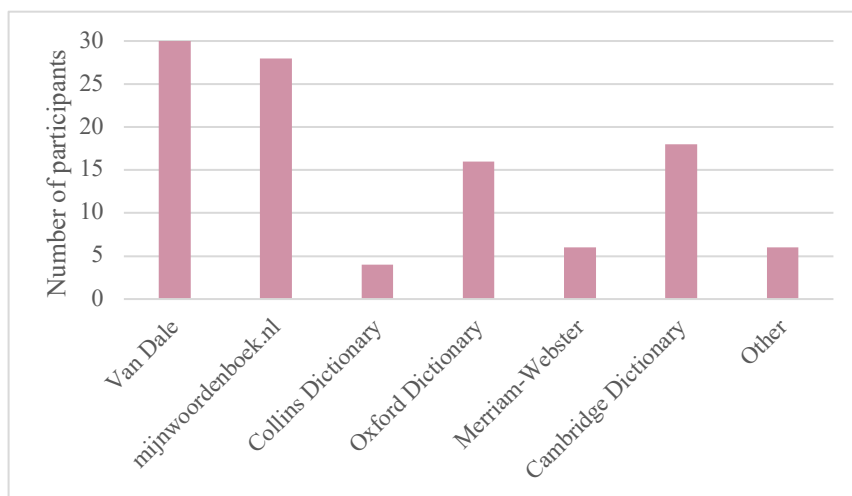


Figure 19: Overview of the dictionaries participants reported having worked with before (total number of participants: 30)

Secondly, the participants were asked what they typically use dictionaries for. The bar chart below illustrates the results (Figure 20). All participants reported that they use dictionaries to look up words. The majority (19 out of 30) also tends to use them to look up expressions. Half of the participants said they use dictionaries to look up phrases. A third of the students reported consulting dictionaries to look up the pronunciation of words. Two students chose the ‘Other’ option because they also use dictionaries to verify spelling or to find synonyms and antonyms.

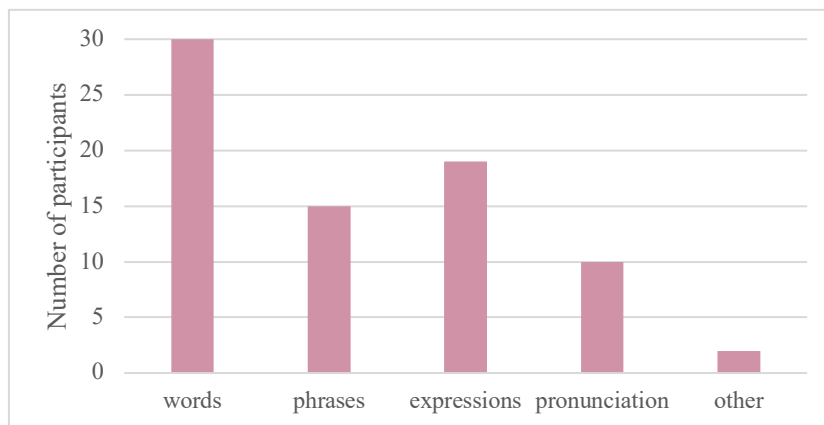


Figure 20: Overview of the types of information participants reported usually looking up in dictionaries (total number of participants: 30)

6.1.2 Machine translation

The second set of questions in this section of the survey gauged the participants’ experiences with machine translation tools. First, they were asked if they had used machine translation tools before. All participants indicated ‘yes’.

Secondly, the participants were asked to specify which tools they had previously used. The analysis revealed that almost all participants (28 out of 30) had used DeepL. Furthermore, the majority had experience with Google Translate (26) and Reverso Translate (24). One participant reported having worked with Papago. Another participant selected the ‘Other’ option to indicate that they had utilised Linguee, a translation dictionary provided by DeepL. The results are presented in the form of a bar chart below (Figure 21).

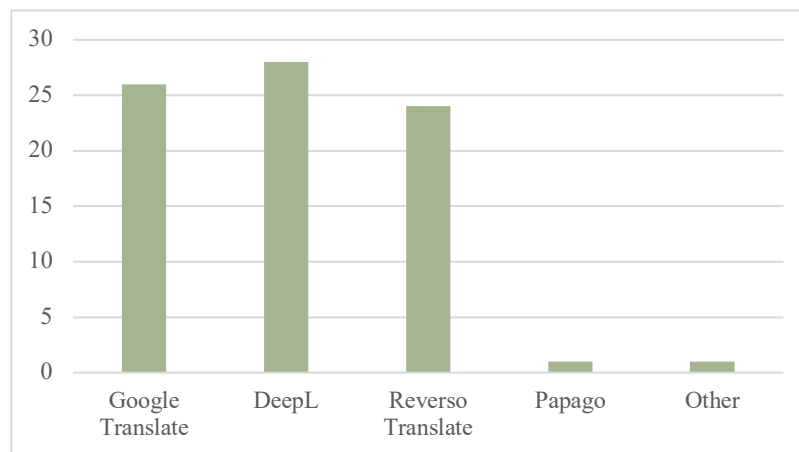


Figure 21: Overview of the machine translation tools participants reported having worked with before (total number of participants: 30)

Next, the participants were asked for which of the four language skills they tend to use machine translation tools. Nearly all participants (29 out of 30) reported using them for written production. A third of the participants said that they use them for reading comprehension. Seven respondents said they used it for oral production, while only one respondent said they used it for listening comprehension.

The third question regarding the participants' general experience with machine translation pertained to the specific goals for which they use it. The majority of respondents (27 out of 30) reported the utilisation of tools such as DeepL and Google Translate for the purpose of vocabulary and expression retrieval. Furthermore, a significant proportion of participants (19 out of 30) reported using these tools to expand their vocabulary. Approximately half of the participants (16) tend to use machine translation to verify the spelling of specific words. Writing grammatically correctly, writing various types of sentences, and checking pronunciation also seemed to be recurring purposes when using machine translation. Lastly, two participants used it for other purposes. They shared that they tend to use these tools to “check if a word works in a certain context, or to find a better synonym”, or to use the output for their translation assignments so they can post-edit it. Figure 22 provides an overview of the functional purposes of machine translation tools as reported by the participants.

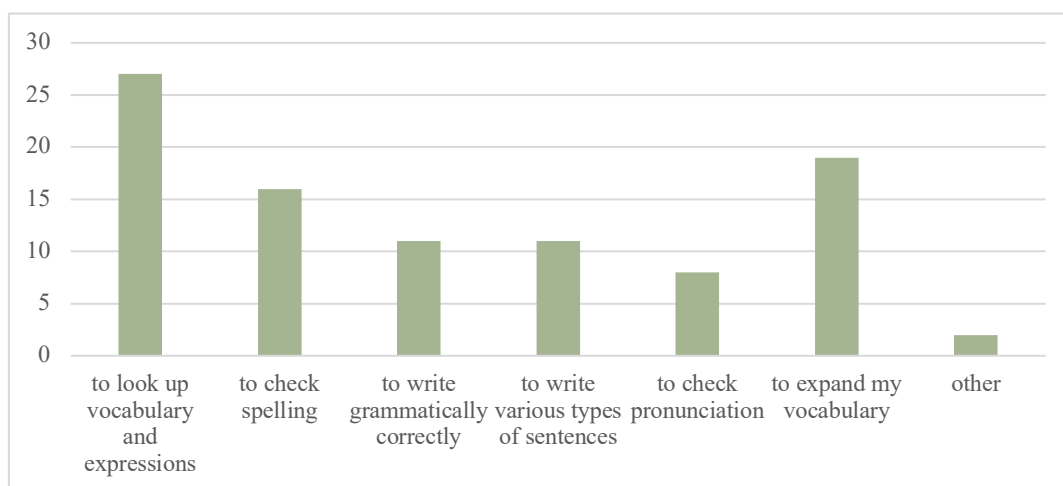


Figure 22: Overview of the functional purposes for which participants use MT tools (total number of participants: 30)

In the final question concerning the general use of machine translation tools, the participants were invited to provide a motivation for their use of machine translation. The predominant reasons for the use of these tools are as follows: firstly, the time-saving properties and convenience of machine translation, and secondly, its ability to review one's work. These options were selected by 26 and 24 participants, respectively. A minority of the participants reported using machine translation tools due to a lack of confidence in their second language(s) (8) or because of the independence it affords them (4). A comparative overview is presented in the subsequent section (6.1.3, Figure 23).

6.1.3 Generative AI

Subsequently, the participants were invited to respond to a similar set of questions regarding their experience with generative AI tools. The first question asked whether they had previously used one of these tools, to which 29 of the 30 participants answered 'yes'.

The participants who had previously used generative AI tools were then requested to specify the tools they had already employed. The vast majority of the participants (28) had previously engaged with ChatGPT. Conversely, Copilot had been used by only seven participants, while QuillBot had been employed by a mere two.

Then, the participants were asked for which language skills they usually employ generative AI. The majority (26 participants) reported using it for written production. Approximately half of them (16) for reading comprehension. Lastly, four participants use it for the purpose of oral production. Noone reported to use AI for listening. Figure 23 compares these answers to the

answers regarding the participants' general experience with MT tools. It shows that both types of tools are predominantly employed for written production. In the second place, participants rely on them for reading comprehension. In addition, a small group of participants uses them for oral production. In each instance, slightly more participants consult MT tools for language assistance than generative AI tools.

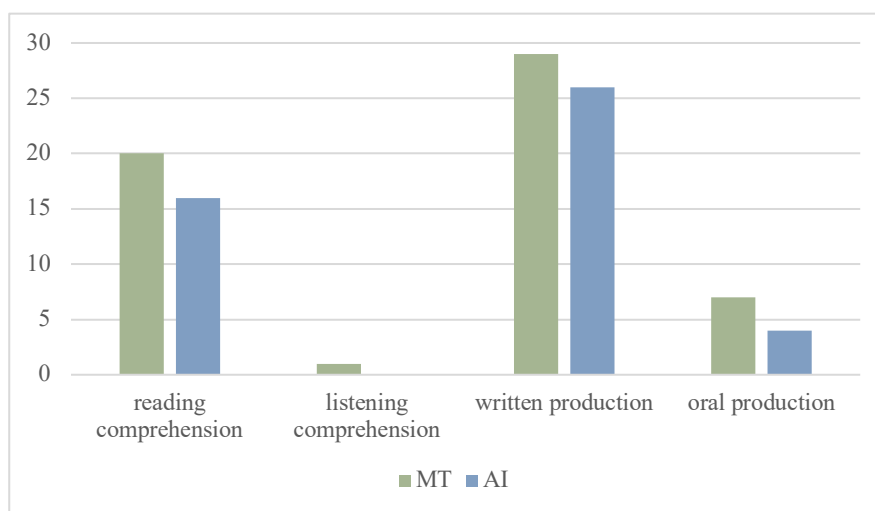


Figure 23: Comparison of MT and generative AI tools in terms of the participants' use for the four language skills (total number of participants: 30 for MT use, 29 regarding the use of generative AI)

After investigating the language skills for which the participants use generative AI, the participants were asked for which functional purposes they generally employ it. The most prevalent answers were the following: “to gather ideas for my argumentation” and “to answer any questions”, reported by 25 and 24 respondents, respectively. A third of the participants uses generative AI tools to generate a draft of their assignment, while a further third uses these tools to receive feedback on their work. Two participants wrote that they also use ChatGPT to (1) “summarise something”, or (2) “to summarise an article, to find useful articles, to translate, to explain a word or sentence”. Figure 24 shows a bar chart of the answers.

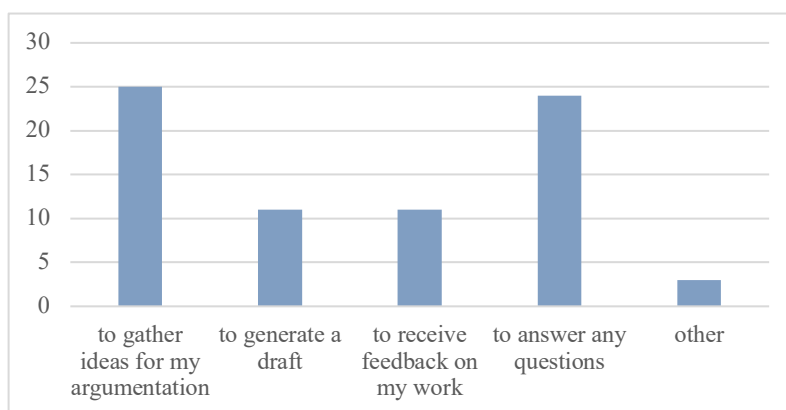


Figure 24: Overview of the functional purposes for which participants use generative AI tools (total number of participants: 29)

Finally, the participants were invited to share why they generally use generative AI tools. In this instance, the argument of the time-saving properties and convenience of the tools was most often selected (26 of the 29 participants). Secondly, participants tend to use these tools because they allow them to check their work. Furthermore, a lack of confidence in the foreign language and independence while studying were each selected by three participants as a motivation to use generative AI. Three participants selected the ‘Other’ option. One participant wrote that they had only used AI on a limited basis, and that it is not a component of their study method. Another participant reported using generative AI tools because they allow them to clarify the topic that they are studying, provide structure, and offer insight. The third participant stated that they utilise ChatGPT to obtain further explanation on specific grammatical topics in both Spanish and English. Figure 25 compares the participants’ motivations to use MT and generative AI tools.

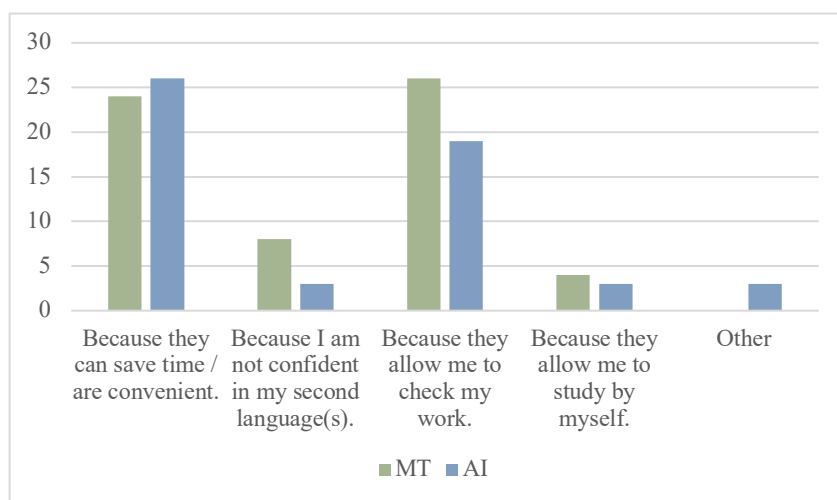


Figure 25: Comparison of MT and generative AI tools in terms of the participants' motivation to use them (total number of participants: 30 for MT use, 29 regarding the use of generative AI)

6.2 Attitudes towards language tools

6.2.1 Perceived enjoyment

After indicating how and why they typically use each type of language tool, the participants were asked whether they usually enjoy using that tool for language learning. The subjects indicated their general enjoyment on a scale of one to five stars. As demonstrated in Figure 26, machine translation tools received the highest average score (4.33 out of 5). All respondents rated them between four and five stars when it comes to enjoyment during their use. Furthermore, dictionaries and generative AI tools were allocated similar average scores – 3.90 and 3.97, respectively. Statistical analysis using a Kruskal-Wallis test showed no significant differences across all three tools ($p = 0.095$). However, the difference between the perceived enjoyment of dictionaries and MT tools is proven to be significant ($p = 0.033$) – in other words, dictionaries were perceived to be less enjoyable to use than MT tools. The dispersion of the results is visualised below (Figure 26).

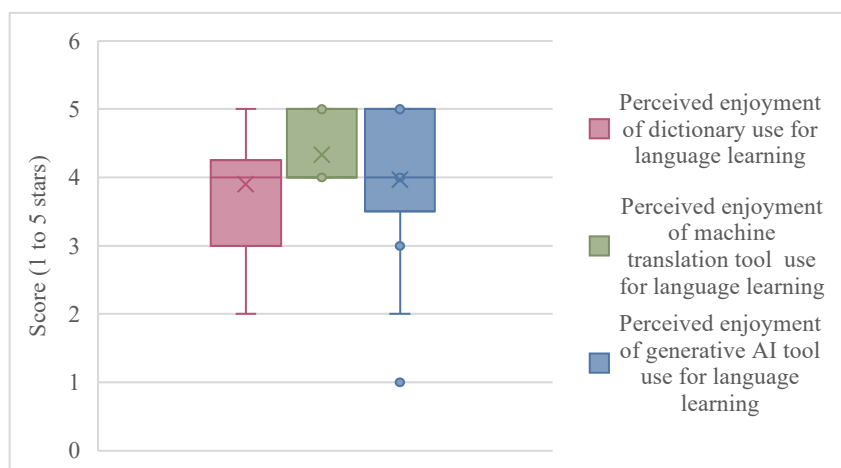


Figure 26: Overview of the general perceived enjoyment of the three types of language tools for language learning, ranked on a scale of 1 to 5 stars

Figure 27 presents the enjoyment scores for each language tool across the three participant groups. This comparison reveals whether a specific group (e.g. the MT-group) felt more positively about their assigned tool (e.g. DeepL) than others. The box plot shows that evaluations of machine translation tools were relatively consistent across all groups, with scores ranging between four and five stars and similar averages. In contrast, enjoyment ratings for dictionaries and generative AI tools varied more noticeably. Dictionaries received their highest average rating from the MT-group (4.29) and their lowest from the AI-group (3.64). As for generative AI tools, they were rated most favourably by the AI-group, with an average score of 4.36 out of 5, while the DI-group gave them the lowest average rating of 3.44. Nonetheless, these differences between language tool groups did not prove to be statistically significant.

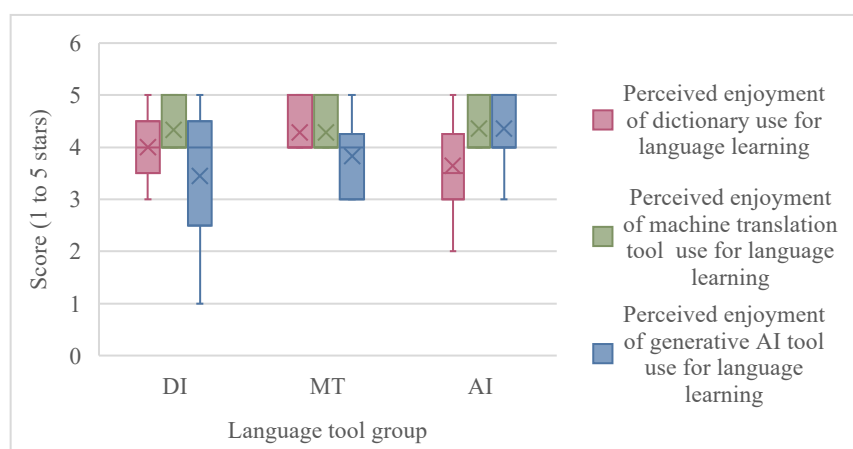


Figure 27: Comparison of the perceived enjoyment of the three types of language tools for language learning across the three groups, ranked on a scale of 1 to 5 stars

6.2.2 Perceived usefulness

In addition to enjoyment, participants were asked to rate the usefulness of each language tool. Figure 28 presents the box plot visualising these scores. The results show that dictionaries were considered the most useful overall, receiving an average score of 4.43 out of 5 stars. Machine translation tools followed with an average score of 4.13, while generative AI tools received the lowest average usefulness rating at 3.97. Although all three tools were generally rated positively, dictionaries and machine translation tools mostly received scores of four or five stars, whereas usefulness scores for generative AI tools are more dispersed. Moreover, the respondents found generative AI tools to be significantly less useful than dictionaries ($p = 0.035$). Differences across all three tools were not found to be significant ($p = 0.091$).

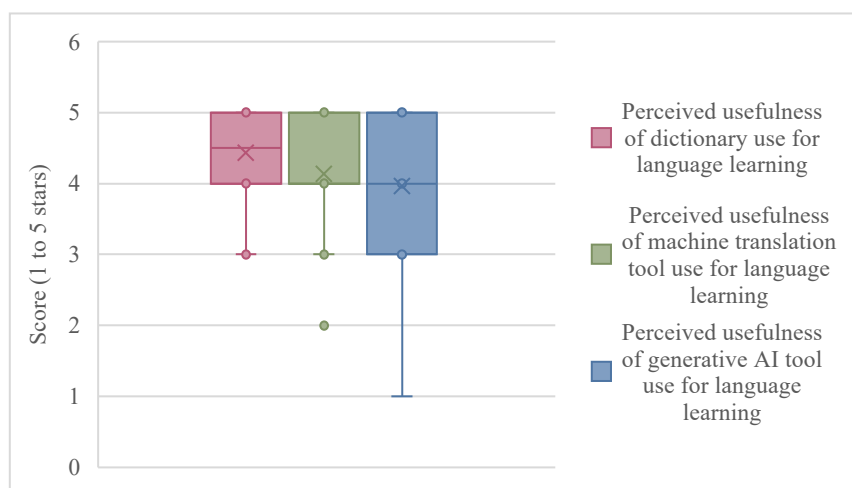


Figure 28: Overview of the general perceived usefulness of the three types of language tools for language learning, ranked on a scale of 1 to 5 stars

As with the enjoyment variable, a box plot was generated to visualise the perceived usefulness of each tool across the three language tool groups (Figure 29). This enables us to determine whether participants evaluated the tool they used for their oral presentation more favourably than the other tools. The results indicate that usefulness scores for both dictionaries and machine translation tools were generally positive across all groups. A particularly notable finding is that the MT-group assigned consistently high usefulness ratings to dictionaries, with almost all participants awarding five stars and only one giving four. Conversely, generative AI tools received lower usefulness ratings from the DI-group, with an average score of 3.44 compared to the overall average of 3.97. Furthermore, the scores for generative AI tools within the DI group displayed greater variability, suggesting a more divided perception of their

usefulness among these participants. Yet, the differences between language groups were not statistically significant.

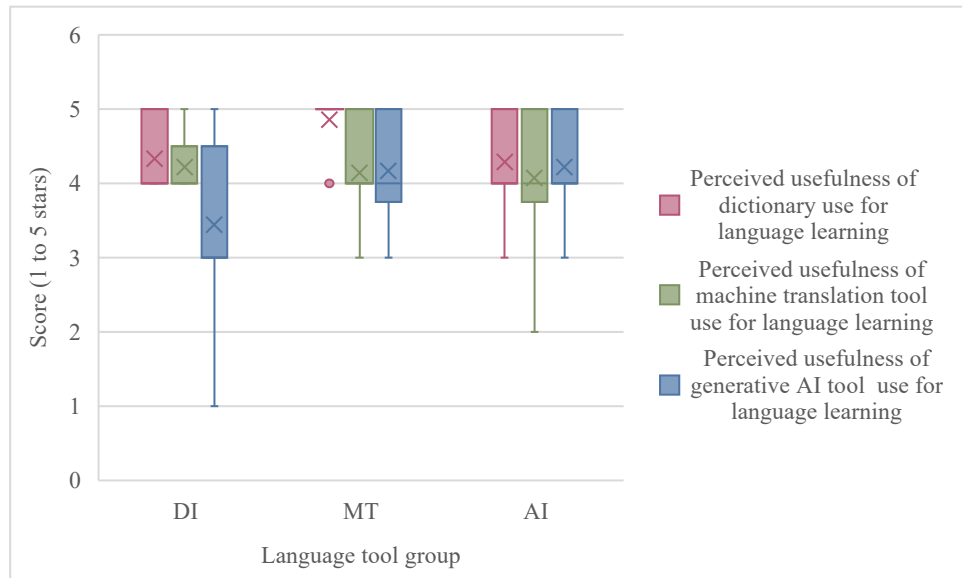


Figure 29: Comparison of the perceived usefulness of the three types of language tools for language learning across the three groups, ranked on a scale of 1 to 5 stars

6.2.3 Perceived advantages and disadvantages

Subsequently, respondents were asked to provide their perceptions of the advantages and disadvantages of each tool. They were provided with six text boxes in which they could enter their answers. The responses can be found in the Appendix. They are categorised in Table 8, alongside quotations from the respondents.

	Advantages	Example quotes	Disadvantages	Example quotes
Dictionaries	Accurate	“Correct information”	Time-consuming	“It takes a bit of time to look up the word”
	Reliable	“100% certainty about the output”, “You can be sure that what is in there is 100% correct.”	Less contextual	“It sometimes gets a little overwhelming to find certain words”, “No words in context, barely any collocations”
	Detailed definitions	“They give clear explanations/definitions of what you look up”	Sometimes outdated	“They are sometimes outdated”
	Good for vocabulary building	“It is easier to learn new vocabulary or to look up vocabulary you don't know”		
Machine translation tools	Fast	“You can translate something you want to say very quickly”	Literal translations	“Machine translations are not always accurate for the translation of the tenses”, “Too literal in their translations”
	Easy to use	“It's fast and easy to use with often different options to choose from”	Grammar mistakes	“It can be wrong and ambiguous. If translating an entire sentence/ part of a sentence, it can be grammatically incorrect”
	Translates sentences	“Able to translate expressions or larger chunks of text”, “You can immediately search a whole sentence or translate a complete text”	Overreliance risk	“You can become a bit lazy because it translates everything for you”
	Helpful for drafts	“Easy to use as draft translations”		
Generative AI tools	Versatile	“You can ask all sorts of questions”, “It not only translates, but also gives alternatives, tips on register etc.”	Can be inaccurate	“Not always correct”, “The information you get can be incorrect”
	Creative	“It gives you more creative answers”	Context errors	“It often makes contextual mistakes or translates too literally”
	Time-saving	“It saves a lot of time”, “Fast way to produce large texts”	Untrustworthy sources	“Not always clear what source is used”, “Can give misinformation”
	Broader support	“You gather information really fast, good to proofread or good to give you a structured basis”	Creativity concerns	“I'm a bit worried about my own creativity”, “It is tempting to let the AI tool do the whole assignment for you”
	Content generation	“They have a lot of information for every question”, “Provides inspiration to start projects”		

Table 8: Summary of the perceived advantages and disadvantages of dictionaries, MT tools, and generative AI tools

6.2.4 Preference

After listing the advantages and disadvantages of the three types of tools, the students ranked the tools according to their preference – from 1 (most preferred) to 3 (least preferred). They provided a ranking of the top three for both speaking and writing. Subsequently, the respondents articulated their rationale behind their choice.

6.2.4.1 Preferred tool for speaking

In preparation for speaking tasks in a foreign language, students most frequently turn to generative AI tools, with 14 out of 30 participants ranking them first. In comparison, machine translation tools received 10 first-place votes, and dictionaries were the top choice for 6 participants. When it came to second-place rankings, machine translation tools were most often ranked second (15 responses), followed by dictionaries (12 responses) and generative AI tools (3 responses). For third-place rankings, generative AI tools were most frequently placed last (13 responses), followed closely by dictionaries (12 responses) and machine translation tools (5 responses). These results are illustrated in Figure 30.

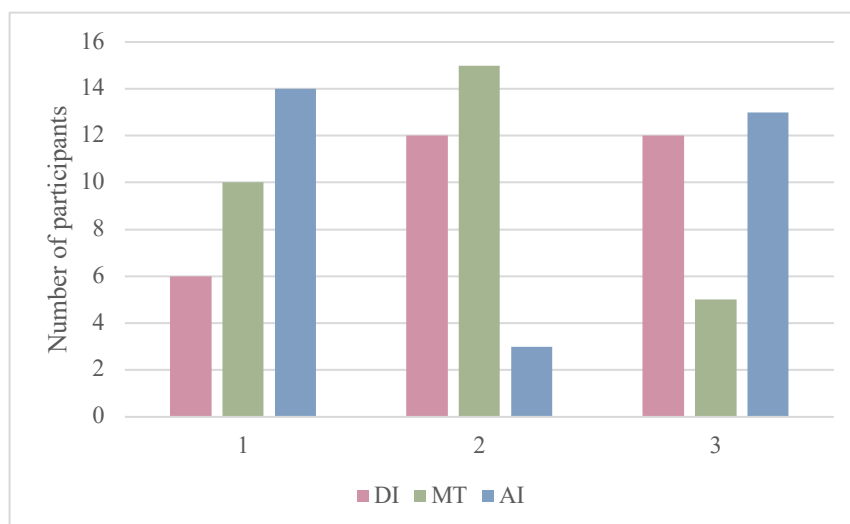


Figure 30: Distribution of the preferred language tool for preparing oral tasks in a foreign language, ranked on a scale of 1 to 3 (total number of participants: 30)

Interestingly, when considering the weighted averages of these rankings, a different image emerges. Calculating the mean scores (where 1 represents the most preferred tool and 3 the least), machine translation tools achieved the highest weighted average of 2.17 out of 3, making

them the overall preferred tool. Generative AI tools followed with a weighted average of 2.03, and dictionaries were rated lowest with a weighted average of 1.80.

These results suggest that while generative AI tools received the most first-place votes, opinions about them were more polarised, as indicated by their relatively high number of third-place rankings. Machine translation tools, on the other hand, appeared to be a consistent middle-ground option that ultimately scored best overall when preferences were averaged. Dictionaries were viewed as less practical for the context of oral task preparation in a foreign language.

6.2.4.2 Preferred tool for writing

When writing in a foreign language, students' preferences for language support tools are slightly different. Machine translation tools were most frequently ranked in first place, with 16 respondents selecting them as their top choice. Dictionaries and generative AI tools each received 7 first-place rankings. In terms of second-place rankings, dictionaries were the most popular second choice with 14 responses, followed by machine translation tools (13 responses), and generative AI tools (3 responses). For third-place rankings, generative AI tools received the highest number of votes (19), indicating that while some students ranked them highly, most considered them the least suitable tool for writing tasks. Dictionaries were ranked last by 9 respondents, and machine translation tools by just 2 participants. The results are visualised in Figure 31.

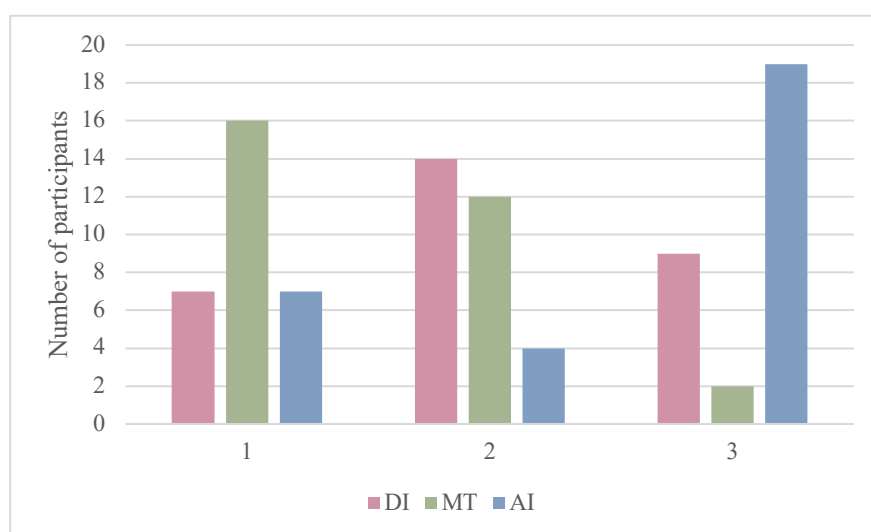


Figure 31: Distribution of the preferred language tool for preparing written tasks in a foreign language, ranked on a scale of 1 to 3 (total number of participants: 30)

When considering the weighted averages of these rankings, the preferences become clearer. Machine translation tools emerged as the overall preferred option for writing tasks, with a weighted average of 2.47 out of 3. This was followed by dictionaries, with a weighted average of 1.93, and lastly generative AI tools, with a weighted average of 1.60.

This indicates that while a number of participants did rate dictionaries and generative AI tools as their preferred option for writing tasks, the overall trend shows a clear preference for machine translation tools. The higher weighted average for MT tools suggests that they were consistently valued across respondents, whereas generative AI tools, despite receiving some first-place rankings, were more frequently ranked lower, resulting in the lowest overall average.

7 DISCUSSION

This study set out to investigate the impact of different language tools – dictionaries, machine translation, and generative AI – on L2 English learners’ oral task performance and their perceptions of these tools. By combining performance-based measures with survey data, we aimed to gain insight into both the linguistic outcomes of tool use and learners’ attitudes towards these technologies. The following section revisits the research questions and discusses the answers derived from the study’s findings.

RQ1: To what extent does the use of dictionaries, a machine translation, or generative AI tool influence L2 learners’ speaking performance in terms of fluency, lexical complexity, and lecturer evaluation?

The analysis of the presentation recordings and transcriptions revealed that the use of machine translation tools resulted in the highest levels of fluency, with the MT-group outperforming both the AI- and DI-groups in terms of total word count, speech rate, and presentation duration. Generative AI tools followed closely, while the dictionary group consistently produced the least fluent presentations. It should be noted that the average speech rate of all presentations was notably lower than the average in British English (Wang, 2021).

Lecturer evaluations reflected the fluency trend, with both the MT- and AI-groups receiving significantly higher scores than the DI-group.

In terms of lexical complexity, generative AI users produced presentations with the highest lexical diversity and density, with statistically significant differences compared to the MT-group for type-token ratio. In contrast, the dictionary group produced language with a higher proportion of high-frequency (K1) words, while MT users used significantly more off-list words. This is in line with findings from previous studies (Kol et al., 2018; Tsai, 2019). However, it could be due to the fact that, proportionally, there were more members of the MT-group discussing the theme of zoos and therefore using names of zoos both in Belgium and abroad, that are considered off-list words. ChatGPT users demonstrated greater use of AWL terms, indicating that their presentations contained a more academic register. Notably, the use of notes during presentations negatively affected fluency while slightly increasing lexical density and diversity.

RQ2: How do learners experience and evaluate the use of these tools when preparing for an L2 speaking task?

Survey data from the first part of the questionnaire showed variation in preparation processes across the three groups. ChatGPT users reported the longest preparation times, though there was considerable dispersion within this group. In contrast to the more language-focused use of dictionaries and DeepL, the AI-group made effective use of ChatGPT's wider functionalities, employing it not only for linguistic support but also for idea generation, answering content-related questions, and drafting their presentation parts.

Acceptance rates for output were similar across the MT- and AI-groups, with most participants accepting between 61 and 80% of the generated content. When rejected, it was often because the output did not align with the intended contextual meaning. In terms of perceived enjoyment and usefulness, dictionaries and ChatGPT received higher average ratings than DeepL. The difference in perceived usefulness was statistically significant, with the MT-group rating their tool lower than the other two. The MT-group's moderate satisfaction ratings correspond with findings from previous research (Rushton, 2020; Sakamoto, 2022). In contrast, the higher satisfaction reported by the DI- and AI-groups is not directly supported by earlier studies.

RQ3: How do learners generally use and experience these tools for language support in their L2?

The second part of the survey provided broader insight into participants' general experience with all three language tools. Usage patterns confirmed that all participants had experience with dictionaries and MT tools, and nearly all had used generative AI tools previously. Across all tools, written production was the primary skill for which these technologies were reported to be applied. Machine translation tools, such as DeepL and Google Translate, were predominantly used to look up vocabulary and expressions – which is in line with previous research regarding MT-supported writing (Kol et al., 2018; Lee, 2020; Tsai, 2019; Chung & Ahn, 2022). Furthermore, they were valued for their time-saving properties, convenience (as in e.g. Chung & Ahn), and ability to review one's work. Generative AI tools, especially ChatGPT, were reported to be commonly used to generate ideas for argumentation and to answer questions, again appreciated for their efficiency and ability to review work.

Attitudinal measures showed that machine translation tools were rated as significantly more enjoyable to use than dictionaries, with generative AI tools receiving slightly more mixed reactions, particularly among DI- and MT-group members. In terms of perceived usefulness, dictionaries and machine translation tools were rated positively overall, though generative AI tools were evaluated less favourably, especially by participants in the DI-group. These positive MT ratings are consistent with prior research (Chung and Ahn, 2022; Tsai, 2019; Rushton, 2020; Sakamoto, 2022), though no studies to date have compared learners' attitudes toward MT, dictionaries, and generative AI within a single study.

Lastly, participants were asked to rank the three tools in terms of their preference for preparing speaking and writing tasks. For speaking assignments, generative AI tools were most frequently ranked first, followed by MT tools and dictionaries. However, when considering the weighted averages of these rankings, a different pattern emerged: machine translation tools were most preferred, followed by generative AI and dictionaries. This suggests that while AI tools attracted some highly positive rankings, MT tools were more consistently preferred overall. One possible explanation is that students were mindful of both the strengths and potential drawbacks of generative AI, whereas they are more familiar and confident using MT tools, which have been available for longer. However, additional research into learner attitudes is required to explore this further.

For writing assignments, machine translation tools were the clear preference, ranked first by a majority of participants and achieving the highest weighted average score. Dictionaries were the second choice, while generative AI tools were least preferred for writing tasks. This pattern highlights a distinct divide between tool preferences for productive language tasks, with machine translation tools occupying a consistently strong position across both modalities.

8 CONCLUSION

This study set out to explore the impact of different language support tools – dictionaries, machine translation, and generative AI – on L2 speaking performance, as well as learners' experiences and preferences regarding these tools. The findings showed that both MT and AI-supported groups outperformed the dictionary group on several speaking performance measures. DeepL users displayed the highest fluency (in terms of duration, word count, and speech rate), while ChatGPT users demonstrated the most complex vocabulary, as evidenced by a higher type-token ratio and greater use of academic vocabulary. Lecturer evaluations also favoured MT and AI-groups over the dictionary group, though it would have been insightful to break these evaluations down by subcategories such as pronunciation, accuracy, content quality, structure, etc.

The survey results revealed clear patterns in tool use and learner attitudes. Machine translation was primarily valued for vocabulary and expression lookups during the writing stage, while AI was particularly helpful for idea generation and answering questions. Interestingly, AI users reported the longest preparation times, though this group also exhibited the most variation in individual preparation durations. In terms of learner attitudes, both the dictionary and ChatGPT groups rated their assigned tool as more enjoyable and useful than the MT group, suggesting that enjoyment and perceived usefulness may not align directly with performance outcomes.

Across the board, participants preferred MT tools for writing tasks and speaking preparation. However, participants expressed slightly more mixed feelings about generative AI tools, particularly those from the dictionary and MT groups.

These findings point to valuable pedagogical implications. A blended approach to tool use, tailored to task type and learner preference, may enhance language learning outcomes. Additionally, explicit training in tool literacy – including how to critically evaluate and appropriately integrate digital tools into language production tasks – would empower learners to make informed decisions and avoid potential pitfalls such as overreliance.

At the same time, this study faced several limitations. The small, unevenly distributed sample, short-term nature of the task, and the artificiality introduced by the recording setting (e.g., wearing microphones, researcher presence) may have influenced the results. Furthermore, the lecturer evaluations were global rather than categorised by aspects such as fluency,

pronunciation, or argument structure, and the analyses did not include measures of grammatical complexity or accuracy.

Future research should build on these findings by conducting longitudinal studies that examine the sustained impact of MT and AI tool use on both writing and speaking performance across different proficiency levels. Incorporating more detailed evaluation criteria and including learner reflections over a longer period could yield deeper insights into how digital tools shape language learning processes, outcomes, and experiences.

References

- Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. Q. (2023). Impact of ChatGPT on Learning Motivation: Teachers and Students' Voices. *Journal of English Studies in Arabia Felix*, 2(1), 41–49. <https://doi.org/10.56540/jesaf.v2i1.51>
- Bulté, B., Housen, A., Pierrard, M., & Van Daele, S. (2008). Investigating lexical proficiency development over time – the case of Dutch-speaking learners of French in Brussels. *Journal of French Language Studies*, 18(3), 277–298. <https://doi.org/10.1017/S0959269508003451>
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching Pronunciation: A Course Book and Reference Guide*. Cambridge University Press. https://www.google.be/books/edition/Teaching_Pronunciation_Hardback_with_Aud/dqgvZq4T4foC?hl=nl&gbpv=1
- Chung, E. S., & Ahn, S. (2022). The effect of using machine translation on linguistic features in L2 writing across proficiency levels and text genres. *Computer Assisted Language Learning*, 35(9), 2239–2264. <https://doi.org/10.1080/09588221.2020.1871029>
- He, Y., & Cardoso, W. (2021). Can online translators and their speech capabilities help English learners improve their pronunciation? In N. Zoghalmi, C. Brudermann, C. Sarré, M. Grosbois, L. Bradley, & S. Thouësny (Eds.), *CALL and professionalisation: Short papers from EUROCALL 2021* (pp. 126–131). Research-publishing.net. <https://doi.org/10.14705/rpnet.2021.54.1320>
- Housen, A., Kuiken, F., & Vedder, I. (Eds.). (2012). *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA* (Vol. 32). John Benjamins Publishing Company. <https://doi.org/10.1075/llt.32>
- Huang, J., & Mizumoto, A. (2024). Examining the effect of generative AI on students' motivation and writing self-efficacy. *Digital Applied Linguistics*, 1. <https://doi.org/10.29140/dal.v1.102324>
- Kellogg, R. T. (1996). A Model of Working Memory in Writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and application* (pp. 57–71). Routledge.
- Khademi, H., & Cardoso, W. (2022). Learning L2 pronunciation with Google Translate. In B. Arnbjörnsdóttir, B. Bédi, L. Bradley, K. Friðriksdóttir, H. Garðarsdóttir, S. Thouësny, & M. J. Whelpton (Eds.), *Intelligent CALL, granular systems and learner data: Short papers from EUROCALL 2022* (pp. 228–233). Research-publishing.net. <https://doi.org/10.14705/rpnet.2022.61.1463>
- Kol, S., Schcolnik, M., & Spector-Cohen, E. (2018). Google Translate in Academic Writing Courses? *The EuroCALL Review*, 26(2), 50. <https://doi.org/10.4995/eurocall.2018.10140>
- Laksana, I. P. Y., Hudiananingsih, P. D., Suciani, K., Virginiya, P. T., & Wahyuni, L. G. E. (2024). *Exploring the Possible Use of Generative Artificial Intelligence in Supporting Students' Speaking Performance*.
- Lambert, C., & Kormos, J. (2014). Complexity, Accuracy, and Fluency in Task-based L2 Research: Toward More Developmentally Based Measures of Second Language

- Acquisition. *Applied Linguistics*, 35(5), 607–614.
<https://doi.org/10.1093/applin/amu047>
- Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, 33(3), 157–175.
<https://doi.org/10.1080/09588221.2018.1553186>
- Lu, X. (2012). The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives. *The Modern Language Journal*, 96(2), 190–208.
<https://doi.org/10.1111/j.1540-4781.2011.01232.1.x>
- Mairano, P., Santiago, F., & Roa, L. C. (2023). Can L2 Pronunciation Be Evaluated without Reference to a Native Model? Pillai Scores for the Intrinsic Evaluation of L2 Vowels. *Languages*, 8(4), 280. <https://doi.org/10.3390/languages8040280>
- Niño, A. (2020). Exploring the use of online machine translation for independent language learning. *Research in Learning Technology*, 2402. <https://doi.org/10.25304/rlt.v28.2402>
- Rushton, A. (2022). Motivating Japanese University EFL Learners to Produce Longer Speaking Turns with Web Based Machine Translation: A Pilot Study. *Kobe Kaisei Jogakuin University Research Bulletin*, 66, 73–81.
- Sakamoto, K. (2022). Japanese university students' reflections on machine translation used as part of an English presentation activity. *人間文化*, 53, 33–45.
- Skehan, P. (1998). *A cognitive approach to language learning* (1. publ., 6. impr). Oxford University Press.
https://www.google.be/books/edition/A_Cognitive_Approach_to_Language_Learnin/Yzdl3pW0Yf4C?hl=nl&gbpv=0
- Stapleton, P. (2021). Using Google Translate as a Tool to Improve L2 Writing: A Case Study of Primary-Level Writing in Hong Kong. *International Journal of Computer-Assisted Language Learning and Teaching*, 11(3), 92–98.
<https://doi.org/10.4018/IJCALLT.2021070106>
- Tang, X. (2025). L2 Writing with AI: Perceptions and Engagement of EFL Learners in China. *English Language Teaching*, 18(2), 68. <https://doi.org/10.5539/elt.v18n2p68>
- Tsai, S.-C. (2019). Using google translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, 32(5–6), 510–526.
<https://doi.org/10.1080/09588221.2018.1527361>
- Van Waes, L., & Leijten, M. (2015). Fluency in Writing: A Multidimensional Perspective on Writing Fluency Applied to L1 and L2. *Computers and Composition*, 38, 79–95.
<https://doi.org/10.1016/j.compcom.2015.09.012>
- Wang, L. (2021). British English-Speaking Speed 2020. *Academic Journal of Humanities & Social Sciences*, 4(5). <https://doi.org/10.25236/AJHSS.2021.040517>
- Yamaoka, K. (2024). ChatGPT's Motivational Effects on Japanese University EFL Learners: A Qualitative Analysis. *International Journal of TESOL Studies*, 6(3), 24–35.
<https://doi.org/10.58304/ijts.20240303>
- Zhang, Z., & Huang, X. (2024). The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon*, 10(3), e25370.
<https://doi.org/10.1016/j.heliyon.2024.e25370>

APPENDIX

A. Assignment sheets

1. DI-group

E2V2 Speaking skills: presentation

In groups of three to four students, you will be asked to give a presentation on a controversial topic of your choice. Please follow these guidelines when preparing for your presentation:

Topic: choose a topic from the list provided earlier.

Duration:

- 15 minutes for a group of 4, 10-12 for a group of 3.
- Each member of the group speaks for about the same amount of time.

Structure:

- Introduction: hook, topic, thesis statement
- Body: arguments
- Conclusion

Presentation aids:

- Beamer (PowerPoint)
- Video/audio
- Blackboard

Tools for preparing this presentation:

- (Online) dictionaries: bilingual and/or explanatory:
 - Bilingual dictionaries, such as Van Dale (as a student at Ghent University, you have full access via **vandale.ugent.be**);
 - Explanatory dictionaries, such as Oxford, Cambridge, Collins, Merriam-Webster.
- For this assignment, do **not** use other tools such as Reverso, ChatGPT, QuillBot, Google Translate or DeepL.
- Of course, you are allowed to use other online sources to gather information on your topic.

Please **time** your preparation! (Approximately, in minutes)

2. MT-group

E2V2 Speaking skills: presentation

In groups of three to four students, you will be asked to give a presentation on a controversial topic of your choice. Please follow these guidelines when preparing for your presentation:

Topic: choose a topic from the list provided earlier.

Duration:

- 15 minutes for a group of 4, 10-12 for a group of 3.
- Each member of the group speaks for about the same amount of time.

Structure:

- Introduction: hook, topic, thesis statement
- Body: arguments
- Conclusion

Tools for preparing this presentation:

- DeepL
- For this assignment, do **not** use any (bilingual and/or explanatory) dictionaries or other tools such as Reverso, ChatGPT, QuillBot or Google Translate.
- Of course, you are allowed to use other online sources to gather information on your topic.

Presentation aids:

- Beamer (PowerPoint)
- Video/audio
- Blackboard

Please **time** your preparation! (Approximately, in minutes)

3. AI-group

E2V2 Speaking skills: presentation

In groups of three to four students, you will be asked to give a presentation on a controversial topic of your choice. Please follow these guidelines when preparing for your presentation:

Topic: choose a topic from the list provided earlier.

Duration:

- 15 minutes for a group of 4, 10-12 for a group of 3.
- Each member of the group speaks for about the same amount of time.

Structure:

- Introduction: hook, topic, thesis statement
- Body: arguments
- Conclusion

Tools for preparing this presentation:

- ChatGPT
- For this assignment, do **not** use any (bilingual and/or explanatory) dictionaries or other tools such as Reverso, QuillBot, DeepL or Google Translate.
- Of course, you are allowed to use other online sources to gather information on your topic.

Presentation aids:

- Beamer (PowerPoint)
- Video/audio
- Blackboard

Please **time** your preparation! (Approximately, in minutes)

B. Survey questions

On which day do you take the E2V2 course?

- ☐ Monday (1)
- ☐ Tuesday (2)
- ☐ Friday (3)

How much time did you spend preparing your part of the presentation? (approximately, in minutes)

Display This Question:

If On which day do you take the E2V2 course? = Monday



Which dictionary (dictionaries) did you use to prepare your part of today's presentation? Multiple answers are possible.

- ☐ Van Dale (1)
- ☐ mijnwoordenboek.nl (2)
- ☐ Collins Dictionary (3)
- ☐ Oxford Dictionary (4)
- ☐ Merriam-Webster (5)
- ☐ Cambridge Dictionary (6)
- ☐ Other: (7) _____

Display This Question:

If On which day do you take the E2V2 course? = Monday

What did you use the dictionaries **for**? Multiple answers are possible. To look up ...

- ☐ words (1)
- ☐ phrases (2)
- ☐ expressions (3)
- ☐ pronunciation (4)
- ☐ other: (5) _____

Display This Question:

If On which day do you take the E2V2 course? = Monday

Overall, did you **enjoy** using dictionaries for preparing this assignment?

1 star: not at all;

5 stars: very
much (1)



Display This Question:

If On which day do you take the E2V2 course? = Monday

Overall, did you think the dictionaries were **useful** for preparing this assignment?

1 star: not useful

at all; 5 stars:
very useful (1)



Display This Question:

If On which day do you take the E2V2 course? = Tuesday

Which type of **input** did you enter into DeepL most of the time? Multiple answers are possible.

- ☐ words (1)
- ☐ phrases (2)
- ☐ clauses (3)
- ☐ sentences (4)
- ☐ paragraphs (5)

Display This Question:

If On which day do you take the E2V2 course? = Tuesday

In **which ways** was DeepL **beneficial** for your preparation? Multiple answers are possible.

- ☐ to look up vocabulary and expressions (1)
- ☐ to check spelling (2)
- ☐ to write grammatically correctly (3)
- ☐ to write various types of sentences (4)
- ☐ to check pronunciation (5)
- ☐ other: (6) _____

Display This Question:

If On which day do you take the E2V2 course? = Tuesday

In which **stage** of your preparation was DeepL especially helpful?

- ☐ When gathering ideas (1)
- ☐ When writing my part of the presentation (2)
- ☐ When practising my part of the presentation (3)

Display This Question:

If On which day do you take the E2V2 course? = Tuesday

How often did you accept DeepL's output? Around ... of the time.

- ☐ 81-100% (1)
- ☐ 61-80% (2)
- ☐ 41-60% (3)
- ☐ 21-40% (4)
- ☐ 0-20% (5)

Display This Question:

If On which day do you take the E2V2 course? = Tuesday

When you did **not** accept DeepL's output, what was (were) the reason(s)? Multiple answers are possible.

- ☐ It was grammatically incorrect (1)
- ☐ The contextual meaning was different from what I had in mind (2)
- ☐ The expressions were not appropriate (3)
- ☐ I was uncertain about the sentences (4)
- ☐ Other: (5) _____

Display This Question:

If On which day do you take the E2V2 course? = Tuesday

Overall, did you **enjoy** using DeepL for preparing this assignment?

1 star: not at all;

5 stars: very
much (1)



Display This Question:

If On which day do you take the E2V2 course? = Tuesday

Overall, did you think DeepL was **useful** for preparing this assignment?

1 star: not useful

at all; 5 stars:
very useful (1)



Display This Question:

If On which day do you take the E2V2 course? = Friday

Which **prompts** did you enter into ChatGPT? Please paste them below

Display This Question:

If On which day do you take the E2V2 course? = Friday

In which ways was ChatGPT **beneficial** for your preparation? Multiple answers are possible.

- ☐ to gather ideas for my argumentation (1)
- ☐ to generate a draft for my part of the presentation (2)
- ☐ to receive feedback on my work (3)
- ☐ to answer any questions (4)
- ☐ other: (5) _____

Display This Question:

If On which day do you take the E2V2 course? = Friday

In which **stage** of your preparation was ChatGPT especially helpful?

- ☐ When gathering ideas (1)
- ☐ When writing my part of the presentation (2)
- ☐ When practising my part of the presentation (3)

Display This Question:

If On which day do you take the E2V2 course? = Friday

How often did you accept ChatGPT's output? Around ... of the time.

- ☐ 81-100% (1)
- ☐ 61-80% (2)
- ☐ 41-60% (3)
- ☐ 21-40% (4)
- ☐ 0-20% (5)

Display This Question:

If On which day do you take the E2V2 course? = Friday

When you did **not** accept ChatGPT's output, what was (were) the reason(s)? Multiple answers are possible.

- ☐ It was grammatically incorrect (1)
- ☐ The contextual meaning was different from what I had in mind (2)
- ☐ The expressions were not appropriate (3)
- ☐ I was uncertain about the sentences (4)
- ☐ The tone was too formal (5)
- ☐ The tone was too informal (6)
- ☐ It would not have sounded credible if I had used the output (7)
- ☐ Other: (8) _____

Display This Question:

If On which day do you take the E2V2 course? = Friday

Overall, did you **enjoy** using ChatGPT for preparing this assignment?

1 star: not at all;

5 stars: very
much (1)



Display This Question:

If On which day do you take the E2V2 course? = Friday

Overall, did you think ChatGPT was **useful** for preparing this assignment?

1 star: not useful

at all; 5 stars:

very useful (1)



Thank you for sharing your experience of this assignment. The next questions assess your general experience with dictionaries, machine translation tools and generative AI tools.

Which dictionary (dictionaries) have you used before? Multiple answers are possible.

- ☐ Van Dale (1)
- ☐ mijnwoordenboek.nl (2)
- ☐ Collins Dictionary (3)
- ☐ Oxford Dictionary (4)
- ☐ Merriam-Webster (5)
- ☐ Cambridge Dictionary (6)
- ☐ Other: (7) _____

What do you usually use the dictionaries **for**? Multiple answers are possible. To look up ...

- ☐ words (1)
- ☐ phrases (2)
- ☐ expressions (3)
- ☐ pronunciation (4)
- ☐ other: (5) _____

Do you usually **enjoy** using dictionaries for language learning?

1 star: not at all;

5 stars: very
much (1)



Do you usually find it **useful** to use dictionaries for language learning?

1 star: not useful

at all; 5 stars:
very useful (1)



Have you used machine translation tools before?

- ☐ Yes (1)
- ☐ No (2)

Display This Question:

If Have you used machine translation tools before? = Yes

Which machine translation tools have you used before? Multiple answers are possible.

- ☐ Google Translate (1)
- ☐ DeepL (2)
- ☐ Reverso Context (the MT tool specifically) (3)
- ☐ Papago (4)
- ☐ Other: (5) _____

Display This Question:

If Have you used machine translation tools before? = Yes

If you use machine translation tools, what types of tasks do you usually use them for? Multiple answers are possible.

- ☐ reading comprehension (1)
- ☐ listening comprehension (2)
- ☐ written production (3)
- ☐ oral production (4)

Display This Question:

If Have you used machine translation tools before? = Yes

What do you usually use machine translation tools **for**? Multiple answers are possible.

- ☐ to look up vocabulary and expressions (1)
- ☐ to check spelling (2)
- ☐ to write grammatically correctly (3)
- ☐ to write various types of sentences (4)
- ☐ to check pronunciation (5)
- ☐ to expand my vocabulary (6)
- ☐ other: (7) _____

Display This Question:

If Have you used machine translation tools before? = Yes

Why do you usually use machine translation tools? Multiple answers are possible.

- ☐ Because they can save time / are convenient. (1)
- ☐ Because I am not confident in my second language(s). (2)
- ☐ Because they allow me to check my work. (3)
- ☐ Because they allow me to study by myself. (4)
- ☐ Other: (5) _____

Display This Question:

If Have you used machine translation tools before? = Yes

Do you usually **enjoy** using machine translation tools for language learning?

1 star: not at all;

5 stars: very
much (1)



Display This Question:

If Have you used machine translation tools before? = Yes

Do you usually find it **useful** to use machine translation tools for language learning?

1 star: not useful

at all; 5 stars:
very useful (1)



Have you used generative AI tools before?

- ☐ Yes (1)
- ☐ No (2)

Display This Question:

If Have you used generative AI tools before? = Yes

Which generative AI tools have you used before? Multiple answers are possible.

- ☐ ChatGPT (1)
- ☐ Quillbot (2)
- ☐ Copilot (3)
- ☐ Other: (4) _____

Display This Question:

If Have you used generative AI tools before? = Yes

If you use generative AI tools, what types of tasks do you usually use them for? Multiple answers are possible.

- ☐ reading comprehension (1)
- ☐ listening comprehension (2)
- ☐ written production (3)
- ☐ oral production (4)

Display This Question:

If Have you used generative AI tools before? = Yes

What do you usually use generative AI tools **for**? Multiple answers are possible.

- ☐ to gather ideas for my argumentation (1)
- ☐ to generate a draft (2)
- ☐ to receive feedback on my work (3)
- ☐ to answer any questions (4)
- ☐ other: (5) _____

Display This Question:

If Have you used generative AI tools before? = Yes

Why do you usually use generative AI tools? Multiple answers are possible.

- ☐ Because they can save time / are convenient. (1)
- ☐ Because I am not confident in my second language(s). (2)
- ☐ Because they allow me to check my work. (3)
- ☐ Because they allow me to study by myself. (4)
- ☐ Other: (5) _____

Display This Question:

If Have you used generative AI tools before? = Yes

Do you usually **enjoy** using generative AI tools for language learning?

1 star: not at all;

5 stars: very
much (1)



Display This Question:

If Have you used generative AI tools before? = Yes

Do you usually find it **useful** to use generative AI tools for language learning?

1 star: not useful

at all; 5 stars:
very useful (1)



Almost there! What are, according to you, the advantages and disadvantages of dictionaries, machine translation, and generative AI tools for language tasks?

○ Dictionaries: advantages (1)

○ Dictionaries: disadvantages (2)

○ Machine translation tools: advantages (3)

○ Machine translation tools: disadvantages (4)

○ Generative AI tools: advantages (5)

○ Generative AI tools: disadvantages (6)

What do you think is the best language tool for preparing **oral tasks** in a foreign language? Please rank the following tools.

_____ Dictionaries (1)

_____ Machine translation tools (2)

_____ Generative AI tools (3)

Please explain why.

What do you think is the best language tool for **writing texts** in a foreign language? Please rank the following tools.

_____ Dictionaries (1)

_____ Machine translation tools (2)

_____ Generative AI tools (3)

Please explain why.

What is your age?

What is your gender?

- ☐ Male (1)
- ☐ Female (2)
- ☐ Non-binary (3)

Which language(s) do you speak at home?

Which languages do you study, apart from English?

- ☐ French (1)
- ☐ German (2)
- ☐ Spanish (3)
- ☐ Italian (4)
- ☐ Russian (5)
- ☐ Turkish (6)
- ☐ Other (7) _____

C. Prompts entered into ChatGPT by the AI-group

-
- AI1.1**
- How do you translate 'gewichtsbeheersing' into english?
 - which title does the minister of health in the US get?
 - what is a friendlier synonym for 'fat'?
 - translate 'breed' from dutch into english
 - can you give a polite description that has the same meaning as 'fat'?
 - how do you translate 'hij is 180 graden gedraaid' (as in he completely changed) into english?
 - is fitness in 'going to the fitness' correct in british english?
 - what is 'gulzig' in english?
 - can one 'adopt a stigma'?
 - do you, in that context, have a synonym for following
 - translate 'bemoeilijken' into english
 - translate 'een rol spelen in' into english
 - what is a more colloquial equivalent of 'moreover'?
 - can you say 'what's more is that ...'? Or does it sound clunky?
 - translate 'de persoon in kwestie' into english
 - give me a synonym for 'difficult'
 - is there one term for 'a person on a weight loss diet'?
-
- AI1.2**
- How to spell broccoli?
 - wat is een goede manier om 'toegankelijkheid van ongezond voedsel' te zeggen in het Engels?
 - Is grocery store British English?
-
- AI1.3**
- Even though genetic traits can make it difficult to lose weight, this does not mean it is impossible to lose weight. In addition, Frayling et al. (2007) identified a gene associated with fat mass and obesity. This gene is also known as the FTO gene. It was found that people with two copies of this gene have a 70% increased risk of obesity. Stunkard et al. (1986) found that when obesity rates in Danish adopted children were compared with those of their biological parents and those of their adoptive parents, there was a strong relationship between the biological cases. There was no relationship between adopted children and their adoptive parents, suggesting genetics plays a role in weight management in adulthood. Like all biological arguments, genetic arguments have a problem with reductionism that you should consider when using studies like this. We also need to consider how generalisable the results are, as one country's results may not reflect another's. → ""It was found that people with two copies of this gene have a 70% increased risk of obesity."" → can you elaborate pls
-

-
- In a broader sense, most of us carry “thrifty genes,” which helped our ancestors survive when food was scarce by storing extra fat. But now, with food so easy to get, these genes can work against us, making it harder to stay slim. → can you elaborate on this as well pls
 - what would be the noun derived from the word satiate?
-

- AI1.4**
- I am doing a group presentation on the stigma surrounding obesity. I'd like to talk about the psychological and social factors such as stress, emotional eating and social injustice. What can i say about these factors?
 - How would you idiomatically translate 'Betreffende psychologische factoren wil ik de focus leggen op stress en emotioneel eten.' from Dutch to English
 - give me synonyms for focus
 - Chronic stress and mental disorders such as depression and anxiety could lead to overeating as a coping mechanism. Could you give me other ways to express 'lead to' in this sentence?
 - what is an easier word for 'predispose'?
 - is emotional eating a category of overeating? or how would you classify this?
 - could you fuse the stress and emotional eating factors together so that some arguments aren't repeated?
 - could you elaborate on the social factors please
 - how do you translate kant-en-klaar
 - how to introduce your last argument in a presentation
-

- AI2.1**
- Make a presentation about 20 minutes long about why or why not should we give money to beggars on the street
 - Give me arguments pro why we should give money to beggars
 - Elaborate each topic in the slides creating an organic text divided into 4 parts that we can use as a guide during our presentation
-

- AI2.2**
- What is the percentage of people that would give money to beggars on the street?
- and**
- AI2.4**
- Make a presentation about 20 minutes long about why or why not should we give money to beggars on the street
 - Give me arguments pro why we should give money to beggars
 - Elaborate each topic in the slides creating an organic text divided into 4 parts that we can use as a guide during our presentation
-

- AI3.1**
- Can you give us the advantages and disadvantages of keeping animals in the zoo?
 - Can you give me some information on the ethical concerns surrounding animals in zoos?
-

- AI3.2**
- Can you give me some information on the ethical concerns surrounding animals in zoos?
 - Are there any alternatives to zoos (to avoid ethical discussions etc.) ?
-

-
- Can you give me some examples of wildlife sanctuaries around the world?
-

AI3.3 • Hi, can you give us the advantages and disadvantages of keeping animals in the zoo? Thank you!

- Can you develop a plan for a presentation on the benefets of zoos?
-

AI3.4 • Can injured or rescued animals be released back into the wild?
 • Can you tell me more about the benefits of zoos, and more specifically about what would happen if the captured animals would be released?

AI4.1 • Can you give me some information on the influence abortion rights have on the mental health of women?
 • Elaborate
 • Could you give some examples of those long-term physical health consequences?
 • Could you give me some examples of physical consequences of being denied an abortion?

AI4.2 • Can you give some arguments that are pro-choice in the discussion about abortion?
 • Can you give some more information about body autonomy?
 • Can you give more information about the circumstances of conception?
 • Can you give more information about children's welfare?
 • Can you give some examples to enumerate arguments in a text?
 • How do you translate 'bovendien' or 'daarnaast'?

AI4.3 • can you give topics for a presentation about abortion
 • topics when you are in favour of abortion
 • give arguments for the consequences of restricting abortion

D. Perceived advantages and disadvantages of dictionaries, machine translation, and generative AI tools

In this section, the answers to the following question can be found: “What are, according to you, the advantages and disadvantages of dictionaries, machine translation, and generative AI tools for language tasks?”

DICTIONARIES

ADVANTAGES	DISADVANTAGES
they show definitions and examples - pronunciation - easy to look up - often shows synonyms	you have to know the spelling of the words you want to look up - the dictionaries do not tell you when you cannot use it (in which context)
they give definitions and usually contexts in which you can use certain words	it takes a bit longer than other tools
It is easier to learn new vocabulary or to look up vocabulary you don't know.	You can't always find sentences with a similar context as the one you need.
Clear information about the word, you immediately also can see the gender and plural of a word. A dictionary is also the most reliable.	it is different to look up a word in its context.
It is clear and you have to do more effort.	It takes a while.
different meanings of a word are given	no translation
multiple meanings are shown	slow
100% certainty about the output	difficult to quickly look up expressions or collocations
Several different meanings all showed at once	Not so easy to use, you need to read a lot of long texts before you find your answer
They allow you to check the spelling for a word, making you more confident about your work because you're using the right words. They also explain words/ groups of words very well.	Unless it's an online version, it can be time-consuming. Some words can also be ambiguous which makes the task harder.
They are very useful to expand my vocabulary.	No disadvantages
finding correct translations and checking whether the spelling of words you use are correct	time-consuming
Correct information	Heavy
wide range of vocabulary	no context and often limited to what is officially recognised (for example, English words in French are used more and more by younger people, but the dictionary will still only provide a translation that is often too long and outdated)
Fast in use.	Limited to single words or certain expressions.
You can look up whether an expression is idiomatic	you do not see a lot of examples in context
You can easily find hundreds of synonyms.	They are restricted to mostly looking up synonyms and definitions.
trustworthy	really hard to find a word
verification	cannot find everything you want

Trustworthiness. Dictionaries like Cambridge and Van Dale are quite reputable.	/
you can quickly look up the meaning of a word, or the spelling	it takes a bit of time to look up the word
It's useful for translations, to find words I don't know how to translate and then choose out of the given options.	It doesn't always have good or many translations.
Surety of how accuracy of meaning	No words in context, barely any collocations
easy access, reliable	you cannot look up sentences
Useful and organised	Heavy to bring to class
In an easy and fast way, you can find useful synonyms, antonyms and idioms.	/
useful aid to sound more idiomatic	not every collocation is in the dictionary
They give clear explanations/definitions of what you look up and also give other extra information such as pronunciation, contextual differences in meaning and if it's formal or informal	It sometimes gets a little overwhelming to find certain words and they are sometimes outdated
You can be sure that what is in there is 100% correct.	I feel like they are not always the most handy to work with.
/	/

MACHINE TRANSLATION TOOLS

ADVANTAGES	DISADVANTAGES
you can check the spelling - has good grammar (mostly)	is not always perfect - very word-for-word or very free translations
easily check expressions, see words in context, grammar is usually somewhat correct	sometimes too literal in their translations
It is quicker than having to look it up in a dictionary.	You can look up sentences, too, but the grammar isn't always correct.
You can immediately search a whole sentence or translate a complete text.	They often make mistakes.
It is fast.	It is not always correct.
can translate a word you do not yet have in your own vocabulary	can give wrong structures of sentences etc.
Context-dependant	No correlation between sentences
easy to use as draft translations	less certainty about accuracy than dictionaries
easy to use, answers are provided immediately, you can see words in the correct context and the sentences (e.g. from Context Reverso) are very useful.	Machine translations are not always accurate for the translation of the tenses.
It can give multiple options and it's fast. It's usually straightforward and it's helpful, which is much appreciated when translating.	It can be wrong and ambiguous. If translating an entire sentence/ part of a sentence, it can be grammatically incorrect.
They're usefull for checking entire sentences.	Very often they are not correct.
finding multiple options for translations for words	some translations are not correct or do not match the right context
Easy to use, quick	Not sure they are always correct

different options available, gives you context and sentences in which the vocabulary is often used	if the vocabulary is not in the database, it will have no available translations or offer synonyms that are not appropriate for the context you are looking for
Able to translate expressions or larger chunks of text.	Takes longer to use than a dictionary.
it saves times, as you can look up short phrases	it is not always correct nor idiomatic
They are easily accessible and help with translations of words or short sentences	Translations are not always accurate, especially with long sentences, and if you still don't understand the sentence, you cannot ask the tool to clarify.
fast	not as trustworthy
fast, easy to access	not always correct
It is quick and easy. You can let it translate entire sentences and even texts.	You have to be critical. The sentences can be translated too literally or they can be grammatically incorrect.
/	you can become a bit lazy because it translates everything for you, but not always with the right semantics
You can translate something you want to say very quickly.	It's often gramatically incorrect or often nonsense.
Collocations and words in context	Not always correct
can look up sentences	less reliable, some constructions may be odd
Time-saving	Not always an accurate translation
It helps me to produce more idiomatic and grammatically correct sentences. It's useful in a broader way, it helps me on textual level (not only phrases or words).	Sometimes I am not sure whether everything is correct. Besides, I am also worrying about my own creativity :)
easier to	not always correct
It's fast and easy to use with often different options to choose from	It often makes contextual mistakes or translates too literally, especially when translating big chunks of text
They are very easy to use and they have a wide use.	I think that machine translation tools don't always give the most idiomatic translations and they also don't take the context into consideration.
You get an answer easily and quickly.	The answer can be wrong.

GENERATIVE AI TOOLS

ADVANTAGES	DISADVANTAGES
quick and straightforward	it gives you all the answers at once, and you don't have to think about the 'why' (why do they use this tense, ...)
specific answer to your specific question	you always have to (or at least should) double check (which can take longer), we cannot always trust it
Nice to find some inspiration.	It is tempting to let the AI tool do the whole assignment for you.
Those are very useful to gather ideas and help you.	I don not find them useful for translating, as they also make a lot of mistakes.

It gives you more creative answers.	The information could be incorrect.
save time	are not 100% correct, even though it looks like they are
Saves time	No reliable source
they can explain difficult concepts, but I do not see their use in language acquisition	you have no idea how the tool came to its results
You can translate faster and you can receive personal feedback on your translations or even suggestions to write more idiomatic sentences.	AI tool still need to improve the translation skills for certain languages.
It's helpful and gives constructive criticism. The tool is always very polite which gives the user a nice experience when using.	It can give you information you weren't looking for/ don't care about. It can also be wrong.
They have a lot of information for every question.	The tools are not authentic sources and can give you false information.
fast answers to questions and help with finding ideas and inspiration	often incorrect information and can be a form of cheating
Easy to use, quick	Can give misinformation
No idea, have no experience with it	From what I can tell, the biggest issue is that it is hard to tell if the source is reliable. For translations, this is a no go in my opinion.
Fast, can give many ideas.	Not fully reliable to get a correct or fitting outcome.
it gathers a lot of ideas and can summarise a text	it is not always correct nor idiomatic
You can ask the tool to clarify, summarise, give content, find articles (content), enhance your text or change the tone. You can also ask about genre specific things, such as APA norms.	The tool doesn't always have access to a paper or article to summarise it even though you can just look it up in browser. Secondly, when you ask to enhance your text, almost always unwanted changes are made so it takes a few queries to achieve more or less the text you want. Finally, when looking up arguments, there's very often some overlap between the arguments given.
fast	not trustworthy, lots of wrong information
fast, easy to access	not always correct
You can enter any prompt and they will help you. They not only translate, but also give alternatives, tips on register etc.	Again, you have to be critical and not blindly trust everything it says.
you gather information really fast, good to proofread or good to give you a structured basis	not always clear what source is used, you also become lazy because it does everything for you. language isn't always that great
It is very useful to ask generative AI tools help with homework, general questions you may have, tips,...	Sometimes it doesn't understand the prompt or it makes mistakes (gives incorrect information)
Specific questions can be answered and or doubts can be cleared	Not always correct
fast way to produce large texts	not reliable
Time-saving and good for inspiration	Not always accurate
It saves a lot of time. AI also provides inspiration to start projects and it helps to improve my foreign languages (correct mistakes, expand vocabulary...).	Again, I'm a bit worried about my own creativity. AI also gives some incorrect facts or sources that don't exist, so it's very crucial to be critical about the information it gives.
saves time when searching information	sometimes erroneous / made up information, no reliable sources

Fast and easy to use, has a lot of different uses and you can always ask for more information if needed	It can make contextual mistakes or it translates too literally, depending on the language you are using it for
You can ask all sorts of questions.	I always feel like I can't totally trust the answers, it is not as trustworthy as a dictionary.
It gives you a lot of info in a short time.	The information you get can be incorrect.

E. Transcriptions

DI1.1

Back in 2018, the online Cambridge Dictionary voted the word nomophobia as their word of the year. Does anyone know what that means? No? It means... It's a noun, and it refers to the fear or worry at the idea of being without your phone, or being unable to use it. And today, we want to talk about the question, does personal tech encourage asocial behaviour, or... and what can be done to address the situation if so? And I believe it does, because personal tech has its advantages and disadvantages. But disadvantages are often things that lead to asocial behaviour in public, because people really start to prioritise their phone over their company, like when you're sitting at the table, and a moment of silence, and everyone just pulls out their phone, and starts chatting with other people, to try not to have awkward silent moments, and the attention starts to get divided between both the group in front of you and your phone, which is not good for the quality of the conversations you have. And I think it's also not good for the development of social skills, because you cannot practice social skills through text or chats, because you really need to be in front of people to learn how to communicate, because via social media, you cannot see their faces and learn the art of non-verbal communication when talking, like little nods with the head or things like that cannot be reproduced through text, and I believe that phones are also an easy solution to the problem of being alone, like waiting on friends, you just pull out your phone and you're never alone, and I believe that some solutions to this problem can be a no-phone zone, like when you hang out with friends, you just put your phone to the side and try not to look at it, and just really talk to your friends, and also teaching people what the effects of... Constantly dividing your attention between everything can be and how bad it can be for your own development. And now I will hand the microphone to XXX.

DI1.2

Okay, so I'm XXX XXX XXX, and we've all seen people in our daily lives on public transportation, in cafes, parks, in class, just sitting on their phones with people next to them that they can talk to, but they don't. And this is a phenomenon that we call thubbing, which is a combination of the words phone and snubbing. So basically the act of rather using your phone and texting other people than talking to the people that are directly in your presence. This thubbing phenomenon can also be perceived thubbing, which is when someone else is using their phone when they are... when you are talking to them, sorry. And this can eventually lead to loneliness and psychological distress, or at least just increase it. And this, when someone else is, let's say XXX is talking to me, and I'm just taking out my phone while she's talking to me, and she's going to start feeling like, what am I doing wrong, you know? And she's just going to go on her phone and not speak to me any further, and it just creates a sort of vicious circle, kind of a ripple effect, and it just never stops because everyone's doing it. Then research shows that in different types of relationships, and more specifically a romantic relationship, this kind of thing can lead to more conflict because your partner can become jealous, or at least more jealous, because they might think that you're texting another person that you are maybe romantically involved with. And in other types of relationships, such as family relationships or friendships, there can be a decrease in trust, and people can feel undervalued, as I've already demonstrated. Because why are you not talking to me, XXX? You can also start feeling a bit out of control, because you can't do anything about the situation without making them also feeling uncomfortable in that particular situation. Now additionally, the use of phones and other

personal devices, such as laptops, maybe in class, can kind of create a barrier at social gatherings, such as family dinners. And research shows that about 89% of Americans, if you ask them now, 89% of them will have taken out their phone while in a conversation with someone else. And 82% of those people have already said that it kind of deteriorated the conversation they were in, which leads to a decreased quality because people don't listen anymore. So, right? And conversations, lastly, also become superficial because people don't care anymore about what you have to say. So you just say the most important things. And yeah. Now I will hand over to XXX.

DI1.3

Okay. So, I'm XXX XXX. As they have said before, the use of personal technology like phones and laptops and other things like that are often used, or often leads to people using that as a digital security blanket, as XXX has said. Older generations also feel like the younger generations are less social because they are constantly on their phone and they prefer texting or messaging over phone calls or in-person meetings. Even though the face-to-face engagement is required for verbal and non-verbal communication to learn what that is. Other than that, there is other negative effects of constantly being on your phone, such as the reduced attention span of younger children, because when they're in public and they just start scrolling, they constantly see small videos of... that take no longer than one minute. So, paying attention to something that does take longer than a minute is harder for them. They also have the possibility to connect with many more people, so it feels like they have many friends, but at the same time, it can lead to a lower level of privacy because many people know what you're doing at what time, and they all know very quickly. But at the same time, it can make you feel more lonely than before because there's not a real life connection, it's only digital. And also, as XXX has said, the increase in use of devices during, for example, family gatherings can lead to more conflicts because there's less quality time between you and the rest of your family, so you know less of each other and you're constantly on your phone. Of course, technology also is a form of convenience. Because, um, you can accomplish many tasks without having to leave your home, so you can do a lot more in less time because you don't need the time to travel from one place to another. Um, so, yes, I think that's it.

DI2.1

Hello everyone, we are XXX XXX, XXX XXX XXX and XXX XXX. Today we will discuss whether taking a gap year is a good idea or not. A gap year is basically a break you take between your studies or during your studies. Many people take it between high school and secondary... university, but it depends. We as a group think that taking a gap year is a good idea, because it has many great advantages. But first I will start by exploring how a gap year has a positive advantage on the motivation of a student and how it can provide academic advantages. So I will start with motivation. It is not unnormal for a student to become burned out after years of continuously studying. A student can even lose their enthusiasm and their drive to study, which leads.... which often leads to poor performances. Therefore, taking a gap year can recharge a student's motivation and it can let a student rediscovers their passions. Because a gap year is not only a break from the academic pressure, it also is... it also gives a student a new sense of purpose. So he or she knows again why they are studying and what they are studying for. So in conclusion, a gap year gives those students a new... a new and fresh perspective on life, but it can also... it can also give the student a new interest, etc. So there are many great advantages, if you ask me. So now I will talk about the academic advantages of a gap year. A student who comes back from a gap year often... often starts his or her education again with a renewed

sense of focus and maturity. It can also give a great clarity of their future and... and the fact why they are studying and what they are studying for. A study also showed that students who come back from a gap year often outperform their peers who went straight into university. And that's because they developed a lot of important life skills such as time management, adaptability, independence and so on. And those skills are very important in life, but also at university and for academic success. Furthermore, some gap year activities like internships or specialized courses can even boost the students' university applications and have advantage for... and have advantage in later life in competitive job fields. Now, XXX will tell you more about it. She will explain how a gap year has... has advantages for the independence of students and how it can give you some extra time to think and to grow.

DI2.2

Taking a gap year gives you the opportunity to discover who you really are, to think about what's your character like, what's your personality like, and what are you good at, what goals do you want to achieve, and that can be useful to make big decisions in life, like what degree do you want to study, what kind of job do you want to have. It also gives you more time to discover some new interests, or to discover some new hobbies, and that's also a way to get to know yourself better, and of course that's also very useful for the future, for your future career. You can decide whether you want to work, or whether you want to go to college or university, and what degree you want to study, because when you have more time to take that important decision, and when you are a bit more at ease, you are more likely to choose the right degree in the first time, without having to switch between degrees later on, because it's not what you expected it to be. You also get the chance to discover new things, new experiences, and in those new experiences you can learn other skills which may be useful for your future, and it may open doors for future careers or future jobs, but also your planning skills will improve, because you only have one year, because usually your gap year is only one year, so you have to organise your own schedule, whether you want to work, whether you want to travel, or to volunteer, so you'll have to make a planning for how long you want to work, and then how long you want to travel, or if you want to find a balance between volunteering and working, or so on, but also on a daily basis, because you don't have the classic structure of going to school, or when you have a job, so you really have to organise your own day. Another benefit of taking a gap year is that you become more independent, not only financially, because you have to find a job for yourself, and you have to gain your own money and use it wisely, but you will also improve your life skills, like cooking skills, doing the laundry, cleaning, which all seem basic skills, but it's very important to become independent. You will also discover new environments, whether that is in another country, or at a new job, you will start by yourself at some point, and you will have to find your way on yourself. And now I'll hand over to XXX.

DI2.3

So, what goes hand in hand with the idea of time is the idea of money. Some people may choose to have a gap year to work and earn some money in order to, for example, pay for university, seeing as higher education can be very expensive. You can also choose to work abroad, where you have the opportunity to learn new skills that can come in handy if you want to keep on studying. And because you're already studying for a year, you have a better idea of what you want to do in life, what you're interested in, and what you want to pursue, which can change the choice of your degree for the better. And you also have already some skills that can come in handy when studying. Working abroad can also, you also have the opportunity to learn a new language or train a language you already know, for example, English. And so that way

you can travel, but also earn money. The only problem with this is that many people have the idea of a gap year as being to travel, but for some people that is just not possible because they don't have the money. Of course, they can work, but working for a whole year is very difficult and they just don't have the opportunity to travel for several months. The so-called rich can travel for almost a whole year, whereas the lower class has to work the whole year, which can lead to a burnout. Sure, working can be useful, but not for a whole year. And this again leads back to the problem of the cost of higher education. Because if things as university weren't as expensive, then the people who don't have a lot of money wouldn't have to save for it, and then they would just be able to use their money for the things they want to do in their gap year. And so ultimately, everyone should have the right to do what they want to do in their gap year. And so in conclusion, a gap year can be good for everybody because you have the time to reflect on what you want to do, but also time to do what you cannot do when studying. You also become independent, and ultimately it's a right that everybody should have. Thank you for listening.

DI3.1

Good afternoon, everybody. My name is XXX XXX, and today we will give you a short introduction to zoos. The existence of zoos has always been a very controversial topic. One can have arguments for both sides, and that's why the three of us will argue a different view, an opposing view. I brought some interesting statistics that can be used to argue either side. For example, in 2020, over US\$ 252.2 million was spent on conservation efforts. So a real effort has been made to preserve endangered species. So that's a very good point of zoos. A rather negative point is the fact that the infant mortality rate in elephants is 65%. This is nearly twice as high as the infant mortality rate in the wild, which would be around 30%. So that might be due to the enclosures that are substantially smaller. But nevertheless, I will argue some contrapoints. There we go. A major argument in favor of the preservance of zoos is the role of education. Annually, 183 million visitors visit zoos in the US in total. For the world, that would be an even bigger number. This increases the knowledge of biodiversity or the knowledge about species and nature in... in total. One cannot love something which you don't know. If you really know what the animals go through, where... how they breathe in real life, this might have an effect on your knowledge of this and your ability to feel for these animals. Another point is the research that is made in zoos. Not many people know this, but there is research in both areas, such as breeding or the interaction among each other or the interaction towards other species. But also for human health, to prevent diseases that are transmittable from animals onto humans. An example of that is the Ebola virus, which happened approximately eight years ago. And the cure was made based on research that was conducted in zoos, because monkeys transmitted on humans and the interaction among monkeys was studied. A third argument is the prevention from extinction. It is known that zoos have prevented at least or more than 48 species around the world. This has become an even bigger challenge because of the climate change, which will endanger even more species, so the role of zoos will become even bigger. A famous example of that is the American bison, around 1,000 of its species were left at the late 19th century, and by now more than 37,000 are present in the wild, so that really worked out well. A last argument is often made against zoos, that it's not well maintained, that there are no strict rules, there are. The AZA stands for the Association for Zoos and Aquariums, this is an American organisation, but we have similar organizations in Europe, prevents the fact that zoos can do whatever they want, so the enclosures are maintained well, the habitats are enriched, so there really is preservation for these habitats. I will now hand it over to XXX, who will give you some contra-arguments to the preservation of zoos.

DI3.2

So as XXX already said, I will be discussing the more negative points towards zoos or not. As you can see on the screen, the first most obvious argument is, in my opinion, that it's really unnatural. Now, taking an animal, taking it out of its natural habitat as you can see in the next one, is just unnatural. It belongs in its own habitat. If you take an animal from Africa and you put it in Belgium, in a makeshift something resembling their natural habitat, it doesn't completely cover all the needs. So that's a very obvious first and second one. Now, a more complicated one would be Expo 58. If there are alarm bells ringing right now, I'll briefly remind you of what Expo 58 was all about. Now, in 1958, the Belgian government thought it was a great idea to take some Congolese people and put them in Brussels on the... in the exposition of 1958 to show off the colonies and colonization, because back then it was a great thing and everybody loved it. And they put them there in a makeshift village to show off what the colonies were about and what the inhabitants of Congo were about. Now, obviously, it was not a great representation, but at that time, it was generally accepted. Everybody thought this is generally okay to do. Everybody loved it. It was educational. It was for leisure, but obviously mostly spectacle. Now, as time went on, more people became critical about the whole Expo 58 thing, rightfully so. And right now, I think everybody condemns it again, rightfully so, because it was just a horrible thing to do. Now, I hope this transformation can kind of happen to zoos as well. From a point where zoos are generally accepted, everybody's like, yeah, this is great, the animals are cared for, it's awesome, to maybe now criticizing a bit more. For example, Boudewijn Seapark has been forced to close by 2037, which is, I think, a turning point for zoos in general. And maybe in 20, 30, 50 years, zoos will have to all close. I personally hope so. Now, I'm not denying that for extinction, for example, it's a great thing, but I... I don't... I think, then, the animals should be in a habitat that's not for show or spectacle. Now, don't be fooled by Pairi Daisa, they're very, very pretty zoos and very, very pretty cages, but they're still cages. It's not because there's no metal poles to look through. It's still a cage. Elephants, for example, have a very, very big cage, again, with water and buildings and everything, but it's still a cage. And I think zoos should be banned to say it like that for these reasons. Now, XXX, I'll hand over to you.

DI3.3

Okay so I'm XXX XXX, and my opinion is a little more nuanced, but similar to XXX, I think that the efforts that zoos put into helping animals should also be taken into account. For example, the ethical rescue, because some... most of the zoos serve as a sanctuary to protect animals from starvation, but also from injuries, deforestation or global warming, and in that case, I think that zoos can really be beneficial to the animals. And to give another example, as Thomas already mentioned, zoos save animals from extinction, but I would like to give another example, that is the example of the Przewalski horses, because in the 1960s, this was the last species of wild horses that was... the last species of wild horses, and they were declared extinct, whereas in zoos, 12 of them were still alive, and thus several zoos worked together on a breeding program, and they were able to increase the... the population to 2,400 by 2018. It doesn't work. Okay. So in that case, I think that zoos may be preserved. However, there is also another aspect of it, namely that in some zoos, animals live in close confinement, and in that case, it can cause serious psychological damage, and I think that this cannot be justified. For example, it can cause mental health issues to the animals, such as OCD for bears, anxiety for giraffes, and clinical depression for clouded leopards, and these mental health issues are caused by the distress of people constantly taking pictures of them, but also due to the changes in diets and activities, and the medical examinations. Furthermore, a study has revealed that the animals

in zoos show abnormal behavior, such as hair-plucking, but also pacing, vomiting, and even self-harm, and I think that in that case the zoos are not a good place to be for animals. So overall, there are many... many arguments in favor of preserving zoos, but on the other hand, the ethical aspect should also be taken into account, so thank you for listening.

MT1.1

So, hello everyone. My name is XXX XXX, and I'll start off by showing you some shocking images. So here you can see a few wild animals that are being held captive in zoos. So you can see a bear chained, a lion behind bars similar to a prison, elephants forced to perform acrobatic tricks, and a few more. So I'll be discussing the mistreatment of animals in zoos today. Many animals are mistreated and malnourished. This is not the case for every zoo, but it is a large number nonetheless. Zoos are mainly used, mainly there for profit. Many zookeepers or zoo owners have said that they will host these exposition days or cool like special days in order to gain more profit. And they will often hold the animals in the second place and the profit in the first place. Then animals that are kept... that are captive also induce psychological harm, such as stress or aggression. They can become very aggressive towards other animals or the zookeepers, or just show aggression in general. And this can also lead to physical harm, whether that is by the zookeepers, so they get beaten or chained as we saw, or they will start to self-harm due to the stress they endure. Then animals can also suffer from zookosis. This is a specific type of psychosis, where they show very atypical or abnormal behaviors such as walking around in circles, swaying back and forth, things that animals really should not be doing. And this is of course due to the confinement, the lack of stimulation, and all these things. These animals are also at risk of dying prematurely. This is partly due to the stress. I will show an article or headline in a bit that will also show the case of this dying prematurely. Then humans interacting with animals can also lead to severe stress for them. For example, if children are constantly touching the glasses of the enclosures, the animals can become very overwhelmed, and this can contribute to the stress that they are already feeling. The enclosures are, again, too small for the animals, which can lead to them fighting other animals in order to gain territory or assert dominance. This has also recently happened in Pairi Daiza. Some zoos, if there is a surplus of animals, some zoos will dispose of them, by either killing them or selling them to unethical... unethical exotic animal dealers, which is of course a very bad practice. And then from all of this, we can conclude that animals deserve to be treated with respect and dignity. They did not choose to be put captive... to be held captive. They cannot consent, they cannot voice their opinion, and when discussing the mistreatment of animals, we have to keep the moral aspect in mind as well. So this, again, is one of the headlines. So the baby panda that dies in Scotland after choking on vomit as nearby fireworks are set off. A panda should not be subjected to loud unnatural noises as fireworks. But this was because it was obviously held in a zoo. And this is a very unsettling title, of course. And then more recently, one bear killed in a fight at Pairi Daiza Zoo, which is of course another very bad thing to happen, and spectators could have been... could have seen that and it would have been very awful for them. And then to put it... to put the actual restricting... like the lack of confinement... no, the actual confinement into perspective. So tigers and lions have about 18,000 times less space in their captive enclosures than what they would normally have where they... if they were to be in the wild. And polar bears have 1,000,000 times less space. This is an incredibly large number and it is very concerning of course wild orcas can swim up to 100 miles a day and dive up to 850 feet. Yet no orca tank measures longer than 220 feet or deeper than 50 feet, so this is more for aquariums. And then black bears average zoo enclosures are 25,000 times smaller than the natural roaming range, which could have also caused the fight that we just saw. Does anyone

remember this specific gorilla? It went very viral in 2016, was a bit of an Internet sensation. This gorilla is Harambe. You may recognise the name or the story. So Harambe was a western lowland gorilla who lived at the Cincinnati Zoo and one day in 2016, a three-year old boy crawled under the fence where a gorilla was kept... was held captive and Harambe then grabbed and violently dragged this boy. And fearing for the boy's life, one of the zoo workers shot and consequently killed Harambe. And this was met with a lot of backlash and controversy in the media due to the... the lethal force that was used. People sought... thought it was very unnecessary to use this and this is once again showcasing how bad it is... how bad the conditions are for these animals and the need for better care. That was all the mistreatment. Now I will give the word to XXX.

MT1.2

Hi, my name is XXX XXX XXX, and I will be talking about the advantages of zoos, because obviously all that XXX just said is true, but we cannot forget that some zoos do get it right, and can be beneficial for both animals as us people as well. Because some zoos have made it their mission to save endangered animal species from well... going extinct by setting up reintroduction programs where they breed endangered species and take care of them, and then when they are ready for it they will release them back into the wild, and by doing so saving the species of course. On top of that, zoologists have done great work at mimicking native habitats of animals, because in the wild animals often face habitat loss, or are being hunted, or threatened by, well, ecological threats, and they do not have to face that in zoos where there is great veterinary care, and where zookeepers have the right foods and medication for the animals ready there. They also have... often have open areas where the animals can roam free, just like they would in the wild, and have the right vegetation and other stuff like that. And to comment on what Marina said earlier about the spaces being much smaller for animals than it would normally be in the wild, for most species they cross these large distances because they are in search of food or they are fleeing from predators, but in zoos they do not have these threats. So that's why for some species it's not really a problem that their spaces are much smaller than they would be in the wild. And... right, it's also beneficial for us people because zoos can operate as educational centres, where we learn about all kinds of animals from all around the world, and they increase our environmental awareness as well. And on top of that, by seeing how vulnerable animals can be and learning about their risks, we also raise our moral obligation to protect animals at risk. And because, sorry. We also, zoo trips can also increase our well-being. Studies have found it can serve as great quality time with family and friends. Even school trips can also be beneficial for our well-being. So zoos can have advantages for the well-being of animals as well as the well-being of us people. And now XXX will talk to you about Belgian zoos versus zoos in other countries.

MT1.3

Good morning, everyone. My name is XXX XXX, and I will be focusing on the comparison between three zoos in Belgium and three foreign zoos. This way, we will explore whether zoos abroad generally provide better or worse conditions than those in our country. So, to start off, I will introduce the three zoos I will be talking about in, yeah, zoos that are located in our country, namely Pairi Daiza, Antwerp Zoo, and Planckendaal Zoo. Pairi Daiza is considered to be one of the most beautiful zoos in Europe. *Wacht hoor*. It has, for instance, large thematic habitats. For instance, the Kingdom of Ganesha, which you can by the way see on the picture on the screen, is meant for Asian elephants. And although the animal enclosures are pretty spacious and immersive, there are still critics that argue that the entertainment-oriented

approach sometimes prioritizes visitors' experience over the animals' needs. So, the people who work, the employees who work in this park, they do try their best to make the habitats as realistic and as beautiful as possible. But they do not always succeed in making the environment safe for the animals. Antwerp Zoo is one of the oldest zoos in the world, and actually the oldest one in Belgium, and it has made some great improvements in recent years when we're talking about animal welfare. But there is still a big issue, and that is the location, because the zoo is located in the heart of the city, which means that space is limited for many species. And then, Plankendael Zoo is best known for its emphasis on breeding programs, especially for endangered species. So, even though they have relatively modern enclosures, some areas still have space and enrichment issues for many species. So, let's now look at the zoos abroad. I have chosen three zoos that are, yeah, they are often praised for their standards. First of all, we have San Diego Zoo, which is located in the USA. Then, we have Singapore Zoo, and then we have the Tiergarten Schönbrunn, I'm very sorry if I pronounced it incorrectly, but it's located in Vienna. So, when we're looking at the enclosures, the San Diego Zoo is exceptionally spacious, and it mimics natural habitats, and it's puts a strong emphasis on enrichment. It also has the advantage of vast land resources and significant funding, which many European zoos lack. Then, Singapore Zoo is world famous for its open concept, which basically means that animals can just move freely around wherever possible without being in, like, small spaces. For instance, orangutans, they can just freely swing in treetops, which you can see on the picture. So, they do care about animal welfare because they let them just walk around, because many animals need to be able to walk around. But it does require a warm climate, which we obviously don't have in Belgium. And then, lastly, the one in Vienna. The zoo in Vienna is actually the oldest operating zoo in the world, but unlike Antwerp Zoo, it has modernized its facilities. Its exhibition combine animal welfare as well as historical surroundings, and they also offer innovative enrichment opportunities for animals. So, how does Belgium compare to these zoos? Belgian zoos, especially Pairi Daiza and Planckendael, are very similar to their... to some international counterparts in terms of conservation and enrichment. However, foreign zoos mostly, they often have advantages in three key areas, and those areas are space, funding, and specialization. When we look at space, for instance, the zoo in San Diego has expansive land, so allowing for larger and more naturalistic enclosures. Then, funding. When we're looking at funding, higher budgets in international zoos enable better research, enrichment, and care opportunities, oh programs, I'm sorry. And then, when we're looking at specialization, for instance Singapore, they have the advantage that they can adopt climate-specific designs, which Belgium, for instance, cannot replicate because our climate, for instance, is completely different than the one in Singapore. Yeah, that was it for my part. Thank you for your attention, and I will now hand over to XXX.

MT1.4

Hello, my name is... yes my name is XXX XXX, and I will elaborate a bit more on the history of zoos, because, of course, they weren't always as we know them right now. So, as XXX said, Tiergarten Schönbrunn is indeed the oldest still operating zoo, but that is not where the concept of zoos starts. It started way earlier, in 1490 before Christ, the first type of zoo was invented. They were called menageries, and they were actually more like private collections for rich people. But animal welfare was, of course, not very important, because they were more a display of wealth. So they were often in very cramped cages, with nothing to do, or, yeah, the overall very bad conditions. But then, in 1752, the Tiergarten Schönbrunn was also a menagerie, but in that year, it became open to the public, and the first official zoo was opened. In the 20th century, there were a lot of big changes, new insights in the way we view animals,

the way we view zoos. So, for example, the animal welfare was really one of the biggest priorities. So, concrete floors were replaced by vast structures, such as grass and stuff. There was a lot of enrichment for the animals, a lot of hiding spots for the animals as well. And that way, they kept evolving to the zoos as we know them right now. It has been proven that animal welfare also contributes to the visitor's experience, because a happy animal means happy visitors. Happy visitors means more visitors. More visitors means more money. More money means better enclosures and better care for the animals. So I think it is true, of course, what XXX said, a lot of animals are being mistreated in the zoos, but we should not forget where we come from. These small enclosures that are now enlarged with a lot of enrichment for the animals, we should not forget that. So, as a conclusion of our presentation, wild animals are sometimes being mistreated, but as Saar said, there are also a lot of benefits from zoos, namely putting endangered species back into the wild. And we can see that Belgium isn't all that bad in comparison with other zoos. So yeah, thank you very much for your attention.

MT2.2

Good morning, I'm XXX XXX. The examples that XXX has just discussed clearly show that euthanasia is about more than just ending someone's life. It's about compassion, it's about wanting to end suffering people are going through, and doing that in a peaceful and dignified manner would allow them to say goodbye to their loved ones without feeling guilt or burden. We would now like to discuss and dissect some arguments that are of new use against euthanasia to explain why we are in favour. Deborah will be discussion... discussing the religious aspect of the debate later on, but I will now talk to you about the autonomy and how we as a society view and deal with death and suffering. One of the main arguments that is often used against euthanasia or assisted suicide is that it is a slippery slope. Some believe that there may be a risk that more vulnerable people would feel like a burden to their environment and feel pressured to consent to euthanasia. And others might say that those who are not in a clear state of mind can't consent. First off, in most countries where euthanasia is allowed, for example in Belgium, there are very strict conditions to be met. You need to be in a clear state of mind, you need to be aware of your request and show that you have carefully considered it. Doctors will evaluate you and your condition to ensure that there is no external pressure and that there is a sound reason for wanting to end your life. It is also important to note that euthanasia is not a right and a doctor can decide not to allow your request if they feel uncomfortable doing so. Secondly, why don't we reverse the argument of a slippery slope? If we force people to either live their suffering or revert to more drastic measures such as suicide, what comes next? So, if we force people to either live their life in suffering or revert to suicide, what comes next? Who are we to decide how someone else chooses to either live or end their life? Now, how many of you would want a loved one to suffer, just so that you wouldn't have to? Why is our pain more important than theirs? Bodily autonomy is the right to choose what happens to your body and your life, and it is a fundamental right that should not be taken away just because we feel unprepared to deal with what comes after. We as a society feel uncomfortable with death and everything that comes with it, and we feel uncomfortable with someone deciding that life just isn't worth living anymore. At the same time, we offer no compassion for the pain that they are in, and we say that someone who commits suicide is a coward, stupid, so much to live for. If only they had held on for their loved ones, they would have spared them all this pain. But what if there's just no other way out? If you're not allowed to choose a peaceful and dignified ending of your life, you're forced to commit suicide and cause all the trauma to all the people around you. Many of those left behind would be spared the guilt, the questions, the what-ifs, if only there had been another option. Compassion,

empathy and personal freedom are the most important aspects of this whole debate, and XXX will now expand on them from a religious point of view.

MT3.2

My name is XXX XXX. One of the reasons why we believe TikTok should be banned is, as has been mentioned before, the presence of dangerous challenges. Now to name a few, we've got for example the blackout challenge. Anyone familiar with the blackout challenge? No? All right, it's essentially a challenge where people are supposed to choke themselves for an extended amount of time until they feel dizzy or high. That is the result of this challenge, and they are supposed to post videos of them doing this on the platform. Now there's a lot of risks to be had with this challenge, and there's already over 20 casualties due to this challenge. Another one is the Benadryl challenge. If you're unfamiliar with Benadryl, it's essentially a drug for allergies. It is also used for sleeping medication, and the challenge here is to use a lot of it. The challenge states that you need to consume 12 to 14 pills of this drug on one day to overdose, to once again get the result of a hallucination. However, this can result in heart seizures, heart failure, seizures, heart failure, and even death. Many children especially were hospitalized during this challenge, and there has been one, at least one, death as well. There's plenty more. You have the one chip challenge, where people consume a dangerously hot chip, or the skull breaker challenge, where a person is supposed to jump up and others kick your legs from underneath you, so you fall on your front or your back, which could also severely hurt people. Now what I'm essentially trying to say here is that there's a lot of them. Many challenges have had a presence on TikTok and were promoted on TikTok, and that is why it should be banned. These challenges spread due to the algorithm of the app, which promotes virality over safety. These clips hold a lot of risks, and people tend to leave more comments and likes on this, and that's why it's promoted even further to become big trends. It's not the first one, it probably won't be the last one. It's been happening for years on TikTok. TikTok doesn't seem to do anything about it, which is why a full ban should be in place, we believe. Especially kids are very vulnerable to this, because they only see the fun part to these challenges and not the risks. I will now hand over to XXX.

MT3.3

Another argument for the banning of TikTok is that TikTok has a tremendously negative effect on the attention span of people, as TikTok reduces the attention... their attention spans. This is especially visible in the brains of children, as their brain is still in the development stage. And there is even a term to refer to this phenomenon of the reduced attention span, and that's the TikTok brain. So neuroscientists have said that the brains of the attention span of children nowadays is... has the five-second... five-second attention span of a goldfish. The TikTok brain, one child obtains this because we scroll every day, they scroll every day on TikTok, and then they encounter only videos ranging from, let's say, 10 to 50 seconds. Some videos are longer, of course. The problem is that each time we see such a video, there is a rush of dopamine in our brain. Dopamine is a hormone which releases a feeling of happiness. However, this feeling only lasts very short, shortly, so our brain keeps seeking new dopamine and new information, which is stimulating. We also see... However, the problem is that children also should work with this attention span and this TikTok brain at school. And as Bais has already said, this could result in poor results at school, because I think we all recall being in a very, very boring math class where we should stay focused, and which requires an attention span to focus for a longer time, up to two hours. The problem is that the attention span of kids nowadays only lasts for, like, keeps seeking new information every minute. However, these classes

require them to stay focused for a longer time, but they lose their focus. And this could also affect, this is also visible in the prestations. Moreover, they struggle with staying focused when they're doing their schoolwork, as this could be very demanding. Another problem is that TikTok plays a major role in the spread of misinformation, as TikTok enables the rapid and uncontrolled spread of misinformation, because you can easily just share your opinion online. And if I agree with this opinion, I can just share it. And people don't always check the reliability of the sources. UNICEF has conducted research on this, and they found that over 75% of children feel unable to judge the reliability of sources online. One reason for this is because they can't distinguish popularity with reliability. So if I were to be a person with only 10 followers online, they might think, okay, that's not a reliable source. But if I were to be an influencer with only over 200 followers online, they might assume, okay, this person is very reliable, and they agree to my opinion. Another reason is because they lack the skills or they haven't mastered the skills to distinguish or verify which is right, because schools nowadays, they simply don't learn children about the perils of social media. However, misinformation can have detrimental effects on their health behavior. For example, during the Corona pandemic, there was this popular conspiracy theory that lemon juice is the solution to the virus, which is, of course, not true. But if I were to be, like I said, a person with over 200 followers and I were to convey this message, children might be influenced by my opinion, and they might try this behavior out, which could, of course, have tremendous effects on their health system. And I'll hand it now over to XXX.

AI1.1

Hi, I am XXX XXX. So, two weeks ago, my brother enraged me. He was arguing that curvy people should be in control of their bodies, that if they are called fat and they feel insulted, nothing stops them from working on their bodies. So he basically claimed, reinforcing the widespread stigma, that curvy people are lazy. He went on to explain how the new Secretary of Health and Human Services of the US, Robert F. Kennedy Jr. , used to be quite full-figured, but that he has turned 180 degrees and is now all healthy because he goes to the gym. We... Well, it doesn't quite work this way. We think the social stigma around obesity is misplaced, because obesity is a complex disease influenced by various factors such as genetics, mental health, the environment. It is not merely fixed by going to the gym, and labelling those affected as lazy or gluttonous, not only ignores their real health challenges, but also adds to their burden. Changing... That was a little spoiler. Losing weight is extremely difficult, as it requires a complete change in lifestyle. Taking on a different diet means buying other, healthier products. And as we know, healthy products are quite expensive, so this can lead to financial problems. And healthy products are not as easy to come by as well, always, because we live in a consumer society with a lot of ultra-processed food. But Lente will elaborate on that. What's more, people on a strict weight-loss diet might find that dining out with their family or going somewhere with friends can pose a problem. As, for example, if their friends want to stop by McDonald's or any other fast-food restaurant, really, they will either not be able to join, because of course, fast-food is not on their diet plan, or they will have to bring their own lunch, which is quite a hassle. Moreover, not only the dieter, but also their family needs to adapt, because for many families it is impossible to cater to every family member's wishes on what's for dinner, for example, so everyone will sometimes have to follow the diet. It is both difficult and a necessity to find a weight-loss plan that fits you as a person, because we all have unique bodies, but most popular weight-loss plans are quite generic, which means that they overlook individual differences and often lead to less sustained success. Research has shown that plans that look at, that are personalized often are more effective in both losing weight as well as maintaining

it, because maintaining weight is also extremely challenging. In fact, half of individuals regain the weight they've lost within two years. And one of the main reasons for this is metabolic... metabolic adaptation, which means that when you start losing weight, your metabolism slows down and creates hunger signals and makes you crave food. This can make it harder to, well, sustain your weight loss, and for many people this stage of the journey is where they lose their motivation and pers... perseverance, so this highlights the complexity of not only losing weight, but also keeping it off long-term. I'll now hand over to XXX.

AI1.2

So, my name is XXX XXX XXX, and I will be demonstrating to you the environmental factors of things that influence, increase the number of obesity in the world, really. And those are the price of unhealthy food options and the availability of them. The first one is the price. Imagine this. I, a student, I live in student housing. That means I get a weekly budget from my parents. It's a tight budget, and I need to buy groceries with it. I come home on a Wednesday evening. I open my fridge, and I see, oh, I forgot to get food. So, I go to the store, and I enter the store. I decide, today is the day I will eat something healthy. I go around the store, and I decide on a typical Belgian dinner. I buy some broccoli, some chicken, and some potatoes. Fairly healthy, nothing special. And I think, I've done a good job. I haven't bought some, you know, fancy broccoli. It's a normal chicken. I think, I will make it, but no. It's almost eight euros for this. I cannot afford that. It's way too expensive to buy food like this every single day. So, what do I have to do? I have to put it back, and I continue my journey through the store, and I come into the freezer aisle, and I see this pizza. This pizza is only three euros. So, what will I do as a student? I will buy the pizza. And this is the reason that some people cannot afford to eat healthy every single day, because as you can see, it's fairly expensive to eat like this every single day. We should not judge people based on the food choices they make, because we don't know what their budget is. They might not be able to afford this. They might be happy that they can afford something to eat. So, that is why. The other one is the availability. Like I said, I am a student who lives in student housing, and it's Friday today, which means I will be going home home later today. When I leave this campus, I know it will be two hours maybe before I'm home. So, when I enter the station with Kato, I will tell Kato, I'm hungry. We have 20 minutes before our train is here. I'm going to get something to eat. The only option in Gent-Sint-Pieters station is vending machines. What is in those vending machines? Crisps, sweets, biscuits. There are no apples in there. So, that is the reason that people choose to eat unhealthy. When they are hungry, the only option they have is sugars and fat. So, I feel we should make healthy foods more available, so people will be more inclined to choose something healthy. I'll hand over to XXX now.

AI1.3

I am XXX XXX. Many people believe that obesity is simply caused by poor self-control, but just as XXX and XXX already mentioned, there are many more factors that influence one's weight. And some of these factors are biological and genetic and completely outside of our own control. Up until now, up until today, scientists have found over 400 genes that are linked to factors such as appetite, metabolism, fat storage and satiation. Let's take the FTO gene, for example, which literally stands for Fat, Mass and Obesity Associated Gene. This gene causes an increase in appetite, reduces satiation and energy, slows down metabolism, and encourages fat storage. When a person receives two copies of a certain variation of this gene, one from their mum and one from their dad, their chances of struggling with weight gain and obesity increase up to 70%. Even more so when a person's parents are both obese, the chance of that

person turning out obese as well is 80% higher than someone who does not face the same challenges. Then there's another phenomenon called thrifty genes. These are genes passed down from our ancestors. For them, these genes were a survival advantage because these genes could help them store fat efficiently. So excess calories were stored as fat, which could later be used as energy during hunting or gathering or farming. They also helped them conserve energy because they had a slower metabolism, so calories were burned slower. And it also made them crave high-calorie foods because high-calorie foods contain lots of fat and sugar, which is more easy to get more energy out of. So as I said, these genes were an advantage for our ancestors, but today they are a disadvantage for us because, just like Lente said, today we have lots of availabilities for food, most often, high-calorie foods. Moreover, in today's modern world, we require far less physical effort since we no longer need to hunt, gather and farm to survive. And for people with thrifty genes, weight loss is more challenging and weight regain also happens more easily. There are also many diseases and health conditions that influence one's weight, but there are simply too many to just pick from and discuss in about 3-4 minutes. I will, however, add that lots of medications also influence one's weight gain, such as antidepressants or antipsychotics. Or diabetes drugs, and many more, and these pills need to be taken daily. So I think that it is important to show some empathy in less of... instead of judging people, and we need to truly understand the medical factors behind obesity and treat it as the medical condition it truly is. And now I'll let XXX talk about the physiological and social factors.

AI1.4

So, my name is XXX XXX XXX. To end with, the psychological and social factors contributing to and perpetuating obesity should be discussed as well. Regarding psychological factors, I would like to direct attention to stress in particular. The correlation between chronic stress and weight gain is that high stress levels cause the body to release cortisol, a hormone that is associated with weight gain, especially around the abdominal area. This biological reaction can help people, can make people more likely to gain weight, no matter their lifestyle choices. Next to that, stress can lead to coping mechanisms such as overeating and turning to high-calorie foods. Foods that are high in fat and in sugar reward the brain's reward, so I'm sorry, they activate the brain's reward system, which makes it more likely for people to overeat them. These foods obviously have addictive qualities to them as well. And on top of that, stress can trigger the release of, the release of dopamine, which is a chemical that increases the cravings for these foods. Then, a subcategory of overeating is emotional eating. This refers to consuming food as a way to deal with, to suppress negative emotions. So feelings like loneliness, like sadness, anxiety, and even boredom, can drive people to seek comfort in foods. This creates a cycle in which... which contributes to... to stress and also, in the end, weight gain. A major source of this stress can actually be weight stigma itself. So people that are obese or overweight often face discrimination in, at work, in healthcare, and also just in social situations in general. And this can add to their already existing stress. So stress should be recognised as both a cause and also a consequence of obesity and stigma really just only worsens this dynamic. Then on to some social factors. I've already briefly mentioned that they can face healthcare, that they can face discrimination in healthcare. What this means is, for example, that healthcare providers can have bias towards people with obesity or people who are overweight. This can lead to them not giving... not giving the patients the medical care they actually require. They can just blame it on their weight and not give them a proper diagnosis on basis, based on that. Another thing is that people with obesity or people who are overweight often avoid seeking medical attention because of fear of judgment, shame, embarrassment, which only worsens their medical conditions. And then to end, cultural and

societal beauty standards. I think everyone in this room has at least once been affected by the pressure of media and societal norms. And the same applies to people with obesity or people who are overweight. How many times have you not seen in the media that thinness equates to success, fame, and most importantly to health? However, health... however, being skinny does not automatically mean that you are living a healthy lifestyle. You could be very skinny, very thin, and still consume fast food almost daily. But you just have these amazing genes and this great metabolism, which makes it harder for you to put on weight. And this reinforcing image just leads to body dysmorphia, body dissatisfaction, shame, and mental health issues in general. In conclusion, obesity doesn't... I'm sorry. In conclusion, obesity doesn't only have to do with personal choices. There's so much more to it. A lot of factors that we have discussed, such as biology, like, such as biology, culture, societal inequalities. And to really make a difference, we should try to stay away from blaming and judging. Focus more on empathy and compassion. And... and most importantly, to push for systemic change and policies that reduce stigma surrounding obesity. Thank you for your attention.

AI2.1

Good morning everyone, good afternoon. Today we're here to discuss an important and thought-provoking question. The question being, why should... should we give money to beggars? It is a dilemma we face on a daily basis, I could say, and opinions on this topic vary widely. Imagine you're walking down the street, and someone approaches you and is asking you for your money. And you are hesitating a bit, because why should you? Should you give the money, or is it better to just walk away? That moment of hesitation is very common, and it's also what inspired us today to talk about this topic. We'll address this issue in three parts. Yasmin will talk about the arguments for giving beggars money, then Greta will talk about the possible consequences, the possible negative consequences of this topic, and Sami will end this presentation by talking about the potential alternatives, and give us a conclusion, and eventually we'll also ask you your thoughts about this opinion. So now I'll give the mic to XXX.

AI2.2

Hi, my name is XXX XXX i, and I will, as XXX mentioned, discuss the reasons as to why we should help beggars and give them money when we come across them in the street. So, I think the main reason as to why most people give money to homeless people is our compassion and our willingness to help others that are in need. It's a very beautiful act of kindness that is almost engraved in us, and I think it's also sort of your ethical duty to help others when you know that you are able to do so. Sorry. I think it is also very reasonable to, to not listen to the stereotypes surrounding homelessness, such as there are a lot of people that say that, you know, homeless people are lazy, homeless people do not want to work, they do not want to improve their situation or change their lives around, but the reality is that not everyone has that opportunity, and while I do believe that it is possible to change your life if you really want to, unfortunately, that is not the case for everyone. Not everyone has that luck to be able to, you know, get off the streets, get a job immediately, and just rebuild their lives. One of the reasons is lack of proper documentation. I'm sure that everyone, every homeless people who is in such a miserable situation, if they had the choice, would just simply apply for a job, but that is just not the case for everyone. Not everyone is allowed to apply for a job, first of all, and another reason as to why you should just directly help people on the street is to offer immediate relief, because by donating, like giving money to charities or donating, you may not be able to help them immediately with their pressing needs, such as food, water, medication. If you feel like they might misuse the money, then you can just ask them to go along with you to the grocery shop,

for example, to, you know, go with them to the pharmacy and buy them some of their necessary needs. Thank you for listening. Now, XXX talk, she will discuss the contra-arguments and the possible negative consequences of giving money to beggars.

AI2.4

Hi everyone, my name is XXX XXX, and I will be... can I get the next? Thank you. It's okay. I will be talking about why we should, based on these contra and pro arguments, why we should think about possible alternatives. Because, as Greta said, just giving them money might not solve the root problem, which could be solved by giving money to local charities and organizations who try to train these people to, first of all, they will give them psychological and physical help and feed them and everything, and their goal is to actually train these people to come back into society as functioning members of the society. So I think that is a very good alternative. So... And then another alternative is that you can advocate for policy changes. By that I mean you can advocate for better housing or better psychological support, which is actually, which are actually examples of the root causes of homelessness and poverty in general. Another way is also by offering non-monetary help. So... even though I don't completely agree with this, because you could give a beggar food or snacks or a gift card or anything but that will not mean that they will get off the street eventually. It's not a long-term way to help, which again doesn't mean that you shouldn't do it. So it's very double, which is basically our conclusion actually, that there should be a middle ground. Sometimes by assessing a specific situation you can know whether you should help or not. So it's also very circumstantial and depending on the person and depending on the situation, which is also difficult again because you don't know, you cannot just know somebody's situation and story. So it's a very difficult subject and we tried to get like a yes or no conclusion, but we could not. So I think a combination of immediate help and indirect long-term help would be a good way to deal with this societal problem. So the middle ground is what we think makes the most sense, or I think. So my question to you guys is how many of you feel guilty when you see beggars on the street for not giving money? Yeah, I do too, but then I have these same thoughts in my head about what will they use the money for? Do I really want to give them my money? Which is selfish in a way because it basically means you don't want to provide immediate help, but of course I understand why people don't give money because there are many reasons as to why you will hesitate as well. So I think this is our conclusion here and I hope you enjoyed our presentation. Thank you for listening.

AI3.1

Hello everyone, I'm XXX XXX, and today we will be talking about pros and cons of animals being kept in zoos. XXX and me will present some counter-arguments, and XXX and XXX will present some arguments in favour. A first concern is the limited space and the impact on the natural behaviour of the animals. Even the best zoos, despite their efforts to keep... to create environments that mimic the natural habitats of the animals, face challenges in fully replicating the dynamics and the complexity of the wild. This can have several consequences on the well-being of the animals. A first consequence is stress. In the wild, animals have freedom to explore, to hunt or to move across large territories. And in captivity, they do not have this freedom, and this can lead to frustration and anxiety. Furthermore, stress can also weaken the immune system, and this can lead to illness. A second consequence is boredom. Their activities are limited to eating and interacting with a few other animals, which is quite boring. And ultimately, they will start to behave unnaturally. Here is an example of a bear killing another bear in Pairi Daiza, which very recently happened. And it implies that animals can get

aggressive towards other animals due to the constraints of their environment. Another concern is the poor conditions in zoos and aquariums, especially when facilities are not properly maintained or inadequately funded. For example, the state of the accommodation of the Boudewijn Sea Park in Bruges, where dolphins are being kept in captivity, is alarming as the facilities appear neglected and with poorly maintained tanks and inadequate environments. When an accommodation looks abandoned, it signals a lack of resources, both in terms of funding and in terms of management. And this implies that they do not only fail in... They do not only fail their duty to care for the... to provide care for the animals, but also to provide the public an accurate representation of animal welfare. Now I'll hand over to XXX who will present some other counter arguments.

AI3.2

So, thank you, XXX. So, hi, I'm XXX XXX, and I would like to talk to you about some other ethical concerns, other than the ones that XXX has already mentioned, of course, as well as some possible alternatives to zoos. So, first of all, there are a lot of ethical concerns surrounding, well, the captivity of animals in zoos, because of, well, a lot of reasons, but I will be discussing two main reasons, other than the ones that XXX has said, of course. So, first of all, there are a lot of breeding programs in zoos, and zoos basically justify that by saying that they are keeping... are conserving the species of animals by using breeding programs, but there has been a lot of discussion about this, because people say that it can have a lot of negative consequences as well, such as inbreeding, especially when they are kept in very small spaces, which XXX has already discussed, of course, and this can lead to a lot of psychological issues as well as physical issues. Also, a lot of people have noticed that the focus has mainly been on the so-called charismatic species, which are basically the species that are focused upon because they are really attractive to the public, such as pandas or tigers and so on, and this means that other species, which are equally endangered but are less marketable, so to speak, such as salamanders and other types of reptiles, are often neglected. Which brings me to my next... my next point, which is the commercialisation of animals. So, by focusing on these charismatic species, it has often been claimed that zoos only care about the profit that they gain by exhibiting these animals, and that they care less about the actual welfare of these animals, and some even claim that animals aren't even seen as sentient beings, but rather as things. So, taking all of these issues into account, a lot of possible alternatives have been presented, and I've listed a few here. So, first of all, modern technology can be used to give people an educational view on animals as well. So, for example, they can... virtual reality and augmented reality can be used to give spectators a interactive experience with animals without actually keeping them in confinement. Also, we can... life... wildlife webcams can be implemented to give the people a look how these animals actually behave in their natural habitat. Basically, with modern technologies, the possibilities are endless. And also, something that can be interesting as well, are wildlife sanctuaries, which are basically these places where animals are taken care of and rehabilitated to go back into the wild, and visiting these can also be a really educational experience for visitors. So, that was it for the disadvantages of zoos, and now I will hand over to XXX and XXX, who will be talking about some possible advantages of zoos.

AI3.3

I'm XXX XXX XXX, and I will be talking about some benefits, because as XXX and XXX told us before, there are quite some disadvantages and dangers related to keeping wild animals in the zoo. But of course, we shouldn't deny the significant benefits of zoos. Firstly, zoos provide the opportunity to educate the people about different species and their habitats. This

might be very important, because many people don't have the chance to see these animals... animals in the wild, in their natural habitats, because they live far, far away mostly. And by learning about these animals, people can develop a greater appreciation for them, and for the importance of protecting their habitats. Zoos also provide the opportunity to learn... to learn the people about conservation efforts and awareness of the environment. Additionally, zoos can also kickstart individuals' interest in biodiversity, which is very crucial for environmental protection. And of course, as we all know, especially children love learning about animals up close in a safe environment. So I think mostly, in fact, education is the biggest advantage of zoos. But of course, there's more than this. Zoos also allow researchers to study animals' behaviour, physiology, breeding patterns and more in a controlled environment, which brought many breakthroughs in animal treatment and behaviour. An example of this breakthrough is, for example, why elephants swing their trunks. And the answer on this question is, they either want to play with the other animal, or they want to threaten them and scare them away. So it's actually up to the other animal to decipher what the elephant means. So these breakthroughs all have been made possible thanks to the zoo's ease of access. But of course, we should note that not all animals behave the same in captivity as they do in the wild. So not all research is possible in zoos. Furthermore, we think that many animals have a better and maybe longer life in zoos than in the wild, because there's a lot more danger in the wild in their natural habitats. But of course, this is only possible and true if they are treated well. Animals need a very good treatment to... It's crucial for their mental and physical health. So proper nutrition, proportioned cages or living places and medical care are essential for their lives to live a fulfilling and healthy life. Now I'll hand over to XXX who will talk about even more benefits.

AI3.4

Hi, I'm XXX, XXX XXX, and I will talk about more benefits. As XXX has mentioned, zoos can be very beneficial for us humans, but also for the animals themselves. And I would like to add to that. The first one, the first addition is that there is a conservation of endangered species. There are many species that risk extinction, and that can be due to habitat loss, poaching, or climate change. And I added a couple of examples below. The Przewalski horse, Arabian oryx, and the California condor all face extinction. And there are breeding programs that can help boost their population numbers, but also ensure genetic diversity. The programs exchange animals all over the world with other zoos. And if we didn't do that, then all our animals would inbreed, and they would have to mate with their brothers and sisters. This would eventually weaken the species. So it's very important that we have that genetic diversity. There are also many injured animals, or those who were rescued, because their habitat is no longer a viable and safe place to live. For them, zoos can be a sanctuary, and a place where they can recover until they can be reintroduced into their natural habitat. Now, the process of reintroduction has been successful in the past, especially for the Arabian oryx and the California condor. But it is also a very complex procedure that needs a lot of factors that... Excuse me. That there are a lot of factors that impact this. The first one is habitat loss. If the animal does not have a safe and viable environment to return to, and this can be due to mostly human activity, such as de... such as deforestation or urbanization, they often need... Excuse me. The second argument is loss of natural instincts. This specifically is a survival instinct. So to find food, to hunt for their food, or to run away from predators. These are issues that animals that especially have been raised in captivity never had to face. So they have never developed this natural instinct. And the last factor is the stress and health issues that it can cause for the animal. Because we are ripping them away from the environment they have known their entire lives, and we are reintroducing them into an arguably more dangerous environment. Therefore, the complex

procedure of reintroduction requires a lot of... a lot of monitoring, but also a significant amount of resources and planning. So if we would be extreme and say every zoo has to close, we would not be able... to reintroduce every animal as well. So that's why we introduce a new solution, or yeah not new, the safari parks, where you can drive with your car into an environment that was simulated and is a closer simulation to their wildlife than our regular zoos. That's why we would like to conclude that zoos are... have benefits, they are beneficial, but only if they meet the right conditions that XXX has mentioned, and if the animals can be happy, safe and healthy. Thank you for listening.

AI4.1

Good afternoon, everybody. My name is XXX XXX, and I grew up with two transgender brothers, which means that they were born female and are now male. And luckily, they had in Belgium the opportunity to make that choice about their body. But unfortunately, this is not the case everywhere in this world. So we would like to talk to you about the freedom of choice of women about their bodies, more specifically abortion. We are strong believers in pro-choice, which means that we believe that a woman should have agency over her own body and is supposed to make decisions for herself. We are going to discuss the positive and negative consequences of being able and not being able to get an abortion, the moral debates surrounding abortion rights, and also the consequences of pro-life restrictions. Both obtaining and being denied an abortion has its influence on the mental health of a woman. Being able to get an abortion mostly has positive consequences. The researchers of the University of California, San Francisco, actually found that most women who got an abortion never, well, are no more likely to experience negative emotions, suicidal thoughts, or mental health symptoms than those who were denied an abortion. And actually, 99% of those women eventually believed that they had made the right decision and felt the common sensation of relief after they had the abortion. Moreover, women who were denied an abortion have more psychological problems, and... such as stress and anxiety symptoms, less life satisfaction, and a lower self-esteem. Of course, it is normal to experience negative emotions. It's absolutely not out of the ordinary when you experience negative emotions after getting a... an abortion. But these emotions are often short-term and are commonly associated with personal factors. Not being able to get an abortion has a lot of negative consequences. Women are faced with financial instability. Also, they have issues, they often have issues with finding secure housing and often have strained relationships. Also being forced to be a parent faces them with the challenge of caregiving, and if they're not ready for this, this can also take a toll on their mental health. Not only does not getting... not being able to get an abortion when you want one take a toll on your mental health, it does also take a toll on your physical health. Depending on the health of the pregnant individual and the circumstances of the pregnancy, this can influence the outcome of the pregnancy. One option is, for example, an increased risk of complications from pregnancy, such as maternal mortality and childbirth risks. But also another outcome is the possibility of already pre-existing health conditions getting worse. Now I would like to let XXX talk about the moral debates of surrounding the abortion rights.

AI4.2

Okay, so as XXX mentioned, I would like to talk to... I am XXX XXX. And as XXX mentioned, I would like to talk to you about the moral debate surrounding abortion. There are obviously two sides to the debate, pro-life and pro-choice. People who are pro-life believe that induced abortions are morally impermissible, and people who are pro-choice believe that they are permissible. These are two sides that frequently clash in protests. However, I'm going to

give some arguments to demonstrate that supporting the pro-choice side is the right thing to do. A first pro-choice argument is that women should have ownership over their bodies. Every woman has the right to decide whether or not she wants to have a baby, and the important choice of having children is a decision that women should be able to make themselves. I'm sorry. And they need to have the right to decide what's best for them, considering their own circumstances and their own lives. Abortion is a deeply personal issue. And respect... Abortion is a deeply personal issue and respecting bodily autonomy affirms the idea that individuals are best equipped to make choices about their own lives and their own bodies. So if a woman is forced to remain pregnant, this violates the autonomy that she should have over her body. Next to that, the consequences under which a pregnancy occurs also plays a big role in deciding whether or not she... to continue the pregnancy. A first example is that pregnancies can occur because of rape or incest. And these circumstances can impact the psychological, social, and emotional well-being of the pregnant person. This actually occurs quite often. According to the National Library of Medicine, an official US health institution, one in twenty women in the United States has experienced a pregnancy from rape, sexual coercion or both in their lifetimes. So if a woman is pregnant and that pregnancy is the result of rape or incest, it can also worsen the survivor's trauma if they are forced to continue that pregnancy. So access to abortion is then seen as a way to restore a means of control to the survivor. A second example is that pregnancies can also be caused because of contraceptive failure. No contraceptive is 100% effective and pregnancies can occur despite precautions. Of course, in these circumstances it is clear that the person or the woman did not want to become pregnant, because she is either too young or she is not in a good place in life to have children. So in these cases it is very important as well that there is an access to abortion, so that it allows these women to address this outcome responsibly. And lastly, not only the women's welfare is important, but the children's welfare is very important as well. Every child should be wanted and loved. Children deserve to be born into families where they are genuinely wanted and where the parents are emotionally, financially, and socially ready to care for them. Unwanted pregnancies may lead to neglect or inadequate care for the children and it can have lasting negative impacts on the mental health and the development of children when they are born in an... in an environment where they feel unwanted. And lastly, a lack of financial resources to support a child can also be a reason why women... why women want an abortion. Because if she is forced to carry the pregnancy to term, it can actually push the family deeper into poverty, and this also has negative outcomes for the children, such as poor education, inadequate healthcare, and housing instability. So ensuring that children are born into families where they are loved, supported and provided for is not only beneficial for individual families, but also for society at large, because poor child welfare results may lead to increased public health costs, higher... higher rates of crime and other societal... societal challenges. And I will now hand over to XXX.

AI4.3

Hi, I'm XXX XXX, and now I would like to talk about the consequences of restricting abortion. In Belgium, we are fortunate to have access to safe and legal abortion. But in many parts of the world, and in several states in the US, governed by Republican policies, abortion bans are applied. These restrictions can have significant and often overlooked consequences. This has an impact on individuals, their families, and the society. It's a common misconception that banning abortion reduces its occurrence. But the truth is that abortion bans do not eliminate the procedure. They only make it more dangerous. According to the World Health Organization, 23,000 women die annually from unsafe abortion, while tens of thousands more suffer from severe health conditions. These figures highlight that staying pregnant often is more dangerous

than having an abortion. Frequently, women turn to self-managed or unsafe abortions using unregulated meds obtained online, or they undergo a procedure without proper medical care. This carries a significant health risk for these women. For example, self-managed abortion can cause incomplete abortion, severe bleeding, or infections, while on the other hand, unsafe abortion are often performed by untrained individuals or with harmful methods. Access to safe or legal abortion is essential to protect women's health and lives. Regions where abortion is legal and accessible, the rates of unsafe abortion drop significantly. For example, in the US, women often travel to neighboring states where abortion is legal. And lastly, economic consequences are profound. Forcing women to carry unwanted pregnancies can lead to financial instability, as many drop out of school or quit their jobs. For us, this is a matter of human rights. Numerous international treaties and constitutions protect the access to safe abortion as a fundamental right. Restricting abortion is a direct attack on women's autonomy, denying the basic right to make decisions about their body. This results in gender inequality and it has an impact on women's health, futures, and economic independence. Instead of imposing bans, we should invest in a solution that prevents unwanted pregnancies and support women's choices. These include comprehensive sex education, accessible healthcare services, and economic opportunities. By focusing on education... education, healthcare, and equality, we can build a society where women are respected, supported, and given the freedom to make decisions about their own body. So, we can conclude that the access to safe and legal abortion is a critical component of healthcare, human rights, and well-being. It is essential that we give women the right to make decisions about their own body. Thank you.

