

# [ Investigating L2 English preposition use by Czech university students: A learner corpus study ]

**Zdeňka Neumanová**

Masaryk University,  
Brno, Czech Republic

**[Abstract]** *This paper examines the use of prepositions in the L2 English speech production of L1 speakers of Czech. The data is sourced from a spoken corpus comprising forty c.15-minute interviews. L2 English preposition use was studied by means of potential occasion analysis, and the results show that prepositions pose a challenge to EFL learners. Careful scrutiny of the data revealed an increasing tendency toward preposition accuracy in speech across proficiency levels A2 to B2. Moreover, it is hypothesized that the participants' incorrect EFL preposition selection is influenced by their L1 knowledge.*

**[Keywords]** *accuracy; EFL; negative transfer; potential occasion analysis; prepositions; speech*

## [1] Introduction

The literature is in agreement that prepositions pose a challenge to EFL learners mainly because there is no easily definable pattern in their use which could be of assistance in making choices in particular contexts. Celce-Murcia & Larsen-Freeman (1999, p. 401) state that prepositions are notoriously difficult to learn and add that long after EFL learners have achieved a high level of proficiency in English, they still struggle with prepositions and that even proficient English speakers exhibit variable performance regarding which prepositions they use for a particular meaning. Parrot (2000, p. 94) supports this argument in *The Cambridge Grammar for English Language Teachers* by defining prepositions as a “major problem” for learners.

Accuracy in the use of English prepositions seems to be an underrepresented area in applied linguistics. Whilst errors in general are widely acknowledged, prepositional accuracy in the speech of EFL learners still poses significant challenges to researchers (for an exception see, for example, Nacey & Graedler, 2015). Previous research on preposition accuracy largely focuses on written production and indicates that preposition usage is one of the most difficult aspects of English grammar for non-native speakers to master, and that preposition errors account for a significant proportion of all EFL grammar errors (Chon et al., 2021; Granger, 2003; Thewissen, 2013). This important, albeit under-researched, area has inspired the present empirical investigation, which aims at casting light on the accuracy of English prepositions in the speech of Czech university-level learners. This paper adds empirical evidence as part of a wider investigation into the use of English prepositions by EFL learners. The focus here is not only on how accurately EFL learners whose first language (L1) is Czech use English prepositions, but also on the accuracy order of L2 English preposition production and possible language transfer.

This paper presents an exploration focusing on preposition use among EFL speakers of Czech at three proficiency levels (A2, B1 and B2), seeking to answer the following research questions:

- 1) How accurately do Czech EFL learners use English prepositions?
- 2) What is the accuracy order of L2 English preposition production at the different L2 proficiency levels?
- 3) What are the most frequent error-prone prepositions?
- 4) Are the most frequent errors Czech learners make caused by L1 transfer?

## [2] Theoretical framework

### [2.1] Prepositions

According to Greenbaum and Quirk (1990, p. 188), prepositions are a closed class of items connecting two units in a sentence and specifying a relationship between them.

Prepositions can be approached on the basis of their meaning, use, or, for example, form. Firstly, regarding form, authors classify prepositions as mono- and polysyllab-

ic (Greenbaum & Quirk, 1990) or simple and complex (Carter & McCarthy, 2006; Biber et al., 1999). Secondly, there are issues of use concerning prepositions. For example, it is possible to study instances where more than one preposition with the same meaning is acceptable in a given context, or to study the discourse in which prepositions occur (see Celce-Murcia & Larsen-Freeman (1999, pp. 414–415)). Lastly, the issues of meaning concerning English prepositions are dealt with frequently in English grammars. Biber, Conrad, & Leech (2002, p. 28) explain that prepositions are linking words which introduce prepositional phrases and specify the relationship between two or more entities that they link, or express other abstract relations. Biber et al. (1999) mention that prepositions have the ambiguous status of having borderline lexical membership while at the same time qualifying as functional words (1999, p. 74), and they add that prepositions can have free and bound meanings. For Biber et al. (1999, p. 74), free prepositions have an independent meaning, and the choice of preposition is not dependent upon any specific words in the context (see example 1 below). In contrast, bound prepositions often have little independent meaning, and the choice of the preposition depends upon some other word – often the preceding noun, verb, or adjective (see examples 2 and 3).

- (1) Late one morning **in** June, **in** the thirty-first year of his life, a message was brought to Michael K as he raked leaves **in** De Waal Park.
- (2) She confided **in** him above all others.
- (3) They've got to be willing to part **with** that bit of money. (Biber et al., 1999, p. 74).

Gráf (2015, p. 116) adds that dependent prepositions “typically introduce an object (wait for somebody) or a nominal complement (an expert in early history, good at languages)” and that they form one semantic unit with the preceding word and their selection is not affected by the word they introduce, while the selection of independent prepositions is affected by the relation (temporal, spatial, or other) to the word they introduce (at work, to work, etc.).

As far as the meaning of prepositions is concerned, Quirk et al. (1985, p. 573) note that “so varied are prepositional meanings that no more than a presentation of the most notable semantic similarities and contrasts can be attempted.”

## [2.2] Potential occasion analysis

Potential occasion analysis involves counting the errors of a given type out of the number of times they could potentially have been made. As a method used to study learner language, it requires not only an error-tagged version of the data, but also a part-of-speech (POS) tagged version of the same data. Therefore, unlike traditional error analysis (hereinafter referred to as EA), which relies on counting the specific types of errors out of the total number of errors in the corpus under investigation (or the total number of words in the examined data), the error-tagged corpus will provide information on the number of preposition errors, while the POS-tagged corpus will reveal the overall number of prepositions, i.e. the total number of times the errors could potentially have been made.

Second Language Acquisition (hereinafter referred to as SLA) stands to gain a lot from what potential occasion analysis has to offer in the study of learners' use of prepositions. The descriptive information it provides serves as a basis for analyzing samples of learner language in their totality. As such, it provides an adequate basis for examining which prepositions are used in/accurately and why some prepositions come to be used more accurately than others.

However, a limitation of potential occasion analysis is that it is target-oriented; that is, it tells us whether Czech EFL learners have acquired the system of prepositions. It sheds little light on the actual processes involved in language learning and acquisition, since it is incapable of describing the interlanguage forms that arise as the learners approximate to target language (TL) norms. For this reason, potential effects of L1 influence will be subjected to closer scrutiny in the final part of this paper.

### [2.3] L1 influence

From the point of view of SLA research, the explanation of errors is probably the most important stage in error analysis. However, it might not always be possible to come to a firm conclusion about the source of an error; partly because specific error sources have not been described with sufficient rigor and partly because not all error sources are unambiguous. According to Ellis and Barkhuizen (2005, p. 66), identifying the source of errors is not an easy task since “many errors are likely to be explicable in terms of multiple rather than single sources”.

To avoid tracing which of the many existing strategies are responsible for particular errors, the central focus of my analysis will be on the potential effects of L1 influence. Jarvis (2000, p. 245) mentions 3 potential effects of L1 influence: (a) intra-L1-group similarities, (b) inter-L1-group differences, (c) first language – interlanguage (L1-IL) performance similarities. Intra-L1-group similarities are found when “learners who speak the same L1 behave in a uniform manner when using the L1s” (2000, p. 254), while inter-L1-group differences are found when “comparable learners of a common L2 who speak different L1s diverge in their IL performance” (2000, p. 254). The third potential effect of L1 influence, L1-IL performance similarities, “is found where learners' use of some L2 feature can be shown to parallel their use of a corresponding L1 feature” (2000, p. 255).

Whilst the first two effects rely on automatic and quantitative approaches, intra-L1-group congruity between Czech learners' L1 and their IL performance does not. Assessing the third effect will require a more qualitative evaluation (see sections 5.2.1, 5.2.2 and 5.2.3)

## [3] Review of research

### [3.1] Preposition errors in learner language

Preposition errors in learner language are a common phenomenon that has been extensively researched in the field of SLA and applied linguistics. The knowledge from numer-

ous studies has revealed linguistic features that have established themselves as “known” EFL features, including non-standard use of prepositions (see, for example, Mauranen, 2012; or Cogo & Dewey, 2012).

Key findings from research on preposition errors indicate that learners use incorrect prepositions, as in *\*It is predicted that the degree to social adaptation will determine...* (Celce-Murcia & Larsen-Freeman, 1999, p. 416). Furthermore, EFL learners seem to underuse the category of prepositions (Granger, 1998, p. 48) on the one hand, and tend to insert redundant prepositions (Cogo & Dewey, 2012, p. 48) on the other.

Moreover, inserting redundant prepositions, as in *“\*we have to study about...”*, is on the list of typical “errors” that most English teachers would consider in urgent need of correction and remediation, even though they appear to be generally unproblematic and present no obstacle to communicative success (Seidlhofer, 2004, p. 220).

### [3.2] Classification of preposition errors in learner corpus research (LCR) studies

In learner corpus research conducted on both spoken and written learner corpora, authors frequently examine the types, frequencies, and patterns of preposition errors.

According to Chodorow, Tetreault & Han (2007, p. 26), a preposition error can be a case of incorrect preposition selection (*\*They arrived to the town*), the use of a preposition in a context where it is prohibited (*\*They came to inside*), or failure to use a preposition in a context where it is obligatory (e.g. *\*He is fond this book*). Gráf (2015, p. 116) draws a distinction between dependent and independent prepositions. Of the total number of 121 errors in the use of independent prepositions in the corpus he examined, 105 errors (87%) involved the prepositions *in*, *at* or *on*. These instances were further broken down into the following groups: (1) on the picture/painting/drawing/portrait (48 instances); (2) in university/school (19 instances); (3) various other instances (38 instances).

Moreover, De Felice & Pulman (2009) point out that the most common prepositions, for example *in*, *of*, and *to*, are also among the most frequent words in the language (2009, p. 512). In their study, they focused on nine high-frequency prepositions to ensure sufficient data: *at*, *by*, *for*, *from*, *in*, *of*, *on*, *to* and *with*. They explain that since these are the most frequent prepositions in English, they expect them to occur with high frequency in learner writing too. In their analysis, they considered just those errors where a preposition is needed, but the one chosen by the student is incorrect.

Finally, in their error tagging manual, Granger et al. (2022, p. 23) mention that researchers who are interested in analyzing all erroneous uses of prepositions should consider both errors involving dependent prepositions (i.e. prepositions that are intrinsically linked to a particular adjective, adverb, noun or verb) and errors involving independent prepositions (i.e. lexical errors affecting single or complex prepositions). The categories should also include cases of omission of a necessary preposition.

Based on the above-mentioned relevant studies and sources, the original categorization of errors into dependent and independent prepositions was broadened, and more

details concerning the use of English prepositions were observed. The study I present here is similar to the articles described above in that it investigates the accuracy of the use of English prepositions. For the purpose of the current paper, preposition errors will also be divided into errors involving dependent prepositions and errors involving independent prepositions, and their patterns and frequencies will be scrutinized carefully.

However, my analysis covers a wider range of details concerning specific preposition errors in conversational discourse in the Czech tertiary education context. Moreover, the accuracy order of L2 English preposition production at three L2 proficiency levels is determined, and the causes of Czech learners' most frequent errors are investigated. Using potential occasion analysis, my research therefore aims to reveal the patterns and frequencies of preposition errors that Czech EFL learners make and to indicate what is typical, as well as what is rare with regard to the accuracy of the use of prepositions in English as a foreign language.

## [4] Data and methods

The corpus scrutinized for the purpose of this study was compiled to investigate grammatical errors in the speech of Czech university learners of English. The corpus comprises forty c.15-minute interviews with 20-year-old ( $SD = 1.4$ ) first-year university students, which were recorded and transcribed between 2020 and 2021.

Since the aim of the current study was to analyze speech produced by students of fields other than English philology, whose degree programmes are not designed to educate graduates with an advanced level of language competence (to learn more about the accuracy of advanced learners of English, see Gráf, 2015), the interviewees were first-year bachelor students from four technical fields of study with different specializations (finance and management, travel and tourism, applied computer science, and engineering for industry).

The learners' level of language competence was assessed using the Oxford Placement Test (OPT), which identified the learners' language proficiency levels and provided a means of placing students at the start of the research. The test has been calibrated against the level system provided by the Common European Framework of Reference for Languages (CEFR). The 40 learners whose speech was analyzed for the purpose of this study were placed into the following categories: A2 (13 learners), B1 (15 learners), and B2 (12 learners).

The development of the oral production elicitation tool reflected the research aims, relevant sources (e.g. Bachman & Palmer, 1996; Ellis & Barkhuizen, 2005; Gass & Selinker, 2008; Gries & Paquot, 2020), and similar relevant studies (Brand & Götz, 2011; Dose-Heidelmayer & Götz, 2016; Gráf, 2015; etc.). The interview was designed to elicit students' authentic speech.

Students were provided with identical, precise and clear instructions, and the interviewer guided the students through prompts to elicit a balanced output. The variety of

the tasks (both monologic production and learner-interviewer spoken interaction) was designed so that the characteristics of the tasks corresponded to the characteristics of TL use and also to match the core features that are among the most essential for learners, including various tenses etc.

The questions were comprehensive, and the open questions allowed the interviewees to produce their answers in their own words. To get the conversation started, students were asked to choose a topic and speak about it for 3 minutes without any interruptions.

*Topic 1: A film you have seen / a book you have read and think is particularly good / bad.*

*Topic 2: A place or a country you have visited and liked.*

The first section of the oral interview was stimulated by the interviewer's question (What topic have you chosen?) when the recording was started.

The second task was based on picture descriptions. The students were gradually given three pictures to describe. The interviewer introduced each picture with the quintessential question *What can you see in the picture?* and suggested several other questions if necessary.

During the last task (free conversation), the students introduced themselves, after which the interviewer posed topical scripted questions (e.g. What can you tell me about your family? Tell me about your school. Do you think English will be useful for you in the future?), which were mostly concerned with familiar topics and with learning English. It should go without saying that some of the benefits of this task type can be re-constructed as weaknesses for lower-level learners. The process was rather free-flowing and indeterminate with talkative and accurate learners, while less talkative (and less accurate) learners were often guided by prescribed questions.

The interviews were transcribed – using oTranscribe – by the author of this paper. The interviews were transcribed using an adjusted version of the Louvain transcription guidelines, which are very clear, systematic, and practical. The guidelines were adapted to suit the purpose of this paper.

After the transcription process, the texts were analyzed and manually annotated for errors by a researcher (L1 Czech) and a British English native speaker (NS). After the NS identified errors, these were manually tagged according to the Louvain Error Tagging Manual Version 2.2 (Granger et al., 2022) by the author of this paper.

As noted above, the same data was error-tagged manually by a British English native speaker and the author of this study. To check for inter-rater reliability, the author of this study coded c. 20% of the data (10 transcribed interviews). The inter-rater reliability measured by Cohen's kappa reached 0.92 for errors coded in the sample, i.e. 96.23% agreement, which might be considered an excellent agreement level (Fleiss et al., 2013).

Following Granger et al. (2022), an important distinction was made between errors (the breaking of a specific linguistic rule) and infelicities (instances of non-erroneous but odd-sounding language) when tagging grammatical errors. The latter were not taken

into account in the current analysis because their annotation might be affected by the personal taste of the annotator.

Moreover, since the current paper uses potential occasion analysis to study the use of English prepositions in the speech of Czech EFL learners, which is a method for examining how accurately learners use linguistic features and involves counting the errors of a given type out of the number of times they could potentially have been made, the corpus under investigation was also part-of-speech (POS) tagged. For POS-tagging, the Constituent Likelihood Automatic Word-tagging System (CLAWS) was used, which has consistently achieved 96–97% accuracy. Precision and recall rates yielded by CLAWS4 were not calculated since it was used to POS-tag c. 100 million words of the original British National Corpus (BNC 1994), and it is considered to be highly accurate.

In my analysis, all of the POS tags listed in CLAWS for this specific part of speech were included. The manual frequency counts were accompanied by a detailed qualitative analysis of the prepositions and their context.

Even though some utterances accounted for two tokens of the same preposition in my data (*A112: ...well in in the picture is girl lying on her bed...*), which might have the effect of increasing the overall number of prepositions used in my data, it reflects the nature of spoken language. Carter and McCarthy (2006, p. 173) argue that repetition should be taken as an effective device for maintaining fluency, while Biber et al. (1999, p. 1,056) state that repetition is presumed to be unplanned or involuntary. All in all, spontaneous speech is, among other things, characterized by the presence of repetition, and it has thus been retained for the purpose of the current analysis. Although repetition was not intended to form a part of the present research, its examination might prove helpful in future research, since the frequency of repetition may affect the interpretation of results, mainly in relation to the overuse of certain prepositions in the speech of EFL learners.

Finally, for the purpose of this paper, preposition errors were divided into two categories: errors involving dependent prepositions and errors involving independent prepositions, each of which was further sub-divided in order to discover the most frequent types of preposition errors Czech EFL learners make. The results of the analysis were then compared with four other corpora of speech produced by EFL learners with different mother tongue backgrounds to help us discover whether specific types of preposition errors are caused by L1 transfer and are therefore L1 group-related, or whether EFL learners in general make errors of these types. Since learners often transfer linguistic features from their native language to the target language (see, for example, Bardovi-Harlig & Sprouse, 2018), which can result in preposition errors in their EFL speech, such a comparison is hoped to provide insights into the patterns of language acquisition that are specific to learners with different mother tongues. Studying the extent of language variation between five corpora and the linguistic preferences of learners from different linguistic backgrounds is hoped to reveal whether specific preposition errors (such as “\*on high school”) are L1 group-related.



## [5] Results

All in all, Czech EFL learners do not seem to be using English prepositions accurately. A closer analysis of all the grammatical errors (i.e. errors that break general rules of English grammar; see Granger et al., 2022) that Czech EFL learners make when speaking English revealed that the most frequent error types counted as average error rates (i.e. numbers of errors per 100 words) are articles (having an extremely high frequency, 30.8%), followed by verb tenses (more than 13% of all errors Czech EFL learners made during the interview) and prepositions. The third most frequent error-prone category is English prepositions – 10.4% of all grammatical errors identified in the analyzed corpus are errors involving independent prepositions. Errors in the use of dependent prepositions account for another 3.9%.

A closer examination of the results of A2, B1, and B2 learners shows that of the 2,766 prepositions in the data, more than 11% were inaccurate – quite a high percentage, which explains their ‘bad reputation’ as being difficult to acquire. On the positive side, the most common prepositions Czech EFL learners use incorrectly are *at*, *in*, and *on* (see section 5.1), which hardly prevent understanding in communication.

In what follows, an overview of the quantitative differences between the scrutinized proficiency levels is presented. However, concerning the proficiency variable, the learners were initially selected on the basis of their mother tongue, and they were later assessed for their proficiency level to obtain information about each learner’s level of proficiency. An unfortunate side effect of not being able to customize the data according to proficiency levels was that I obtained different amounts of data in the groups at the six levels of proficiency. This mainly precluded valid comparisons from the proficiency perspective, including all six levels of proficiency. For this reason, only A2, B1, and B2 proficiency levels became central to this paper.

The results are given with the proficiency levels grouped – see Table 1.

**Table 1** Potential occasion analysis

	A2	B1	B2	Total
Prepositions	820	938	1008	2766
Errors	119	115	81	315
Error rate	14.5	12.2	8.0	11.3
Correct uses	85.4%	87.7%	91.9%	88.6%

**Note:** Prepositions = the frequency with which prepositions were used in the speech of the learners in this study.

Error rate = a normed score of errors per 100 words.

Table 1 reveals an increasing tendency toward preposition accuracy in speech across proficiency levels. The target groups that form the main focus of this study, i.e. the A2 to B2 learners, exhibit an increasing tendency with regard to the accuracy of English prepositions in their speech. The analysis reveals that more proficient learners of English seem to err less frequently in the use of English prepositions. In total, at B2, there are

81 preposition errors out of 1008 occurrences of prepositions, meaning an error rate of 8.1 (i.e. 91.9% of the correct uses). Contrarily, 85.4% of the correct uses at A2 indicated that these learners struggle with English prepositions far more frequently than their more proficient counterparts.

All in all, the results show that of the 2,766 prepositions in the data, more than 11% were inaccurate – quite a high percentage, which supports the argument that prepositions are difficult to acquire.

## [5.1] Classification of errors in the use of prepositions

Preposition errors in the corpus under investigation were tagged as LSPR (independent prepositions), or X\*PR (dependent prepositions). LSPR (*AS03: I study <LSPR corr="on"> at </LSPR> Saturday and Sunday because <LSPR corr="on"> in </LSPR> Monday I all day from eight to seven I am <LSPR corr="at"> in </LSPR> the school...*) covers lexical errors affecting single or complex prepositions. Errors in the use of dependent prepositions were further divided into XADJPR (adjectives used with the wrong dependent preposition: *I hate English because it's hard <XADJPR corr="for"> to </XADJPR> me I like I would like to speak English... (FIN2)*), XADVPR (adverbs used with the wrong dependent preposition: *...is a good town but it is too far <XADVPR corr="from"> of </XADVPR> my home... (TT13)*), XNPR (nouns used with the wrong dependent preposition), and XVPR (verbs used with an erroneous, missing, or redundant preposition: *...if we asked we asked them <XVPR corr="for"> on </XVPR> something they have help... (FIN4)*). The category X\*PR covers all errors involving prepositions that are intrinsically linked to a particular adjective, adverb, noun, or verb. The category also includes cases of the omission of a necessary preposition (Granger et al., 2022, p. 22).

The most frequent errors Czech EFL learners make are errors in the use of independent prepositions, i.e. lexical errors affecting single or complex prepositions (LSPR). The error types across proficiency levels are set out in Table 2.

**Table 2** Error types and their frequencies across proficiency levels.

	A2	B1	B2	A1-C2 (Total)	%
<b>LSPR</b>	89	80	59	228	72.4%
<b>XVPR</b>	21	27	19	67	21.3%
<b>XADJPR</b>	4	6	2	12	3.8%
<b>XNPR</b>	5	1	0	6	1.9%
<b>XADVPR</b>	0	1	1	2	0.6%
Total	119	117	81	315	100%

**Note:** LSPR (independent prepositions) = lexical errors affecting single or complex prepositions.  
X\*PR (dependent prepositions) = all errors involving prepositions that are intrinsically linked to a particular adjective, adverb, noun or verb.

As shown in Table 2, independent prepositions are by far the most frequent type of error Czech EFL learners make, followed less frequently by verbs used with an erroneous, missing, or redundant preposition.

### [5.1.1] Independent prepositions

As mentioned in 1.2.2., Chodorow, Tetreault & Han (2007, p. 26) distinguish between three types of errors in the use of prepositions: (1) incorrect preposition selection; (2) use of a preposition in a context where it is prohibited; and (3) failure to use a preposition in a context where it is obligatory. Their distinction was adopted in the current paper. After all the prepositions were identified in the corpus, they were divided into these three groups – see the examples below.

**Table 3** Error types and their frequencies.

Error type	Raw freq.	%	Example
Incorrect preposition	153	67.1%	... I study <LSPR corr=>on>> at </LSPR> Saturday and Sunday because <LSPR corr=>on>> in </LSPR> Monday I all day from eight to seven I am <LSPR corr=>at>> in </LSPR> the school... (AS03)
In a prohibited context	40	17.5%	... week and I visit my grandparents <LSPR corr=>0>> in </LSPR> two times a month... (FM03)
Failure to use where obligatory	35	15.4%	...It's about one guy he was in he was <LSPR corr="in the military"> military </LSPR> ... (AS02)

**Note:** Note that only the targeted error is corrected in the examples, while the other errors are left uncorrected.

Unsurprisingly, the most frequent errors in the use of prepositions are those where learners select an incorrect preposition. The most common type of error in preposition usage is selecting an inappropriate preposition for a given context. This is followed by errors where prepositions are used in contexts where they are not allowed. The least error-prone category is the omission of prepositions in contexts where they are required. Moreover, by far the most error-prone prepositions are *at*, *on*, and *in*, followed by the slightly less frequent *to* and *for* – see the following table.

**Table 4** Error-prone prepositions

Error type	Raw freq.	%	Example
In	69	35.6%	... goes here <LSPR corr="to"> in </LSPR> the university here. (TT03)
On	56	28.9%	... Yeah different subjects than <LSPR corr=>at>> on </LSPR> secondary school... (AS03)
At	31	15.9%	... things when we are <LSPR corr=>on>> at at </LSPR> vacation... (TT12)
To	21	10.8%	...but it's really really much information <LSPR corr=>for>> to </LSPR> me right now... (AS05)
For	17	8.8%	... I visited England <LSPR corr=>two times>> for two times </LSPR> when I was in secondary school... (TT07)

Nevertheless, given that *in* is among the most frequent words in the language and *on*, *at*, *to*, and *for* are among the most frequent prepositions in English (De Felice & Pulman, 2009, p. 512), we can deduce that the frequent occurrence of errors in the utilization of these prepositions can be attributed to the multitude of opportunities where they could have potentially been misused.

Even though the only goal of communication should not be to get the message across – for fear of ‘pidginizing’ the English language (Mukherjee & Rohrbach, 2006, p. 210) – the majority of preposition errors include prepositions *in*, *on*, and *at*, which hardly prevent understanding in communication and do not seem to have the potential to hinder understanding. Their inaccurate use should not be considered a serious problem.

### [5.1.2] Dependent prepositions

Verbs used with an erroneous, missing, or redundant preposition are the most frequent error type in the category of errors in dependent prepositions, in which learners frequently substitute a dependent preposition for an incorrect one (AIII: *I just get <XVPR corr="onto"> into </XVPR> the train buy a ticket...*). Less frequently, Czech EFL learners fail to use a dependent preposition in a context where it is obligatory (AS03: *I like reading and <XVPR corr="listening to music"> listening music </XVPR>*) and the least error-prone category in the use of dependent prepositions seems to be their use in a context where they are prohibited (AS05: *what I do I <XVPR corr="want"> want to </XVPR> more money*).

Moreover, the most error-prone verbs seem to be *go* and *look* - see the following examples: (4) and (5).

(4) ...she looks young she is <XVPR corr="looking at"> looking in </XVPR> the phone... AS03

(5) ...US friends <XVPR corr="went on a"> go to </XVPR> road trip to the... AS05

Out of the 70 occurrences of verbs used with an erroneous, missing, or redundant preposition, both *go* and *look* were used 14 times with an incorrect, missing, or redundant preposition.

Finally, the examples (6) and (7) demonstrate the second most frequent category of dependent prepositions: adjectives used with the wrong dependent preposition.

(6) ...Irish English is a little bit different <XADJPR corr="than"> between </XADJPR> American English or England English... (FIN2)

(7) ... I like I would like to speak English but it's <XADJPR corr="hard for me"> hard to me </XADJPR> to learn English... (FIN2)

While understanding the frequencies and types of learner errors is crucial, it is equally important to investigate possible sources of these errors. S. P. Corder (1981) suggested specific basic steps by which error analysts, who are interested in errors which occur in one's non-native language, may operate when conducting an error analysis. One of these steps is an explanation of errors EFL learners make.

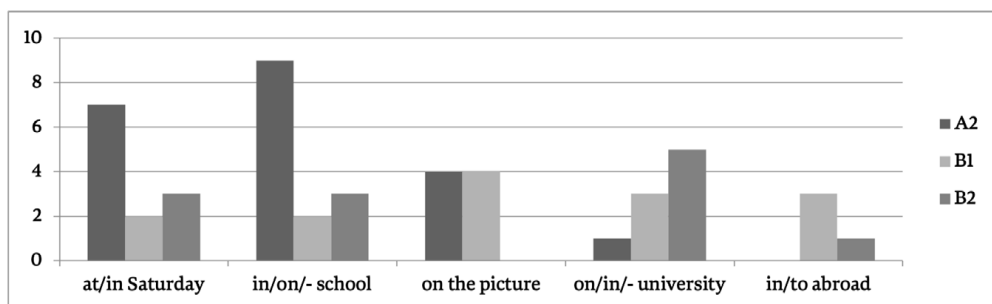
## [5.2] L1 transfer

This part of this paper focuses on potential transfer effects on the Czech EFL learners' use of English prepositions. All phrases with erroneous prepositions were first extracted from the Czech corpus under investigation with the help of WordSmith Tools 8.0 (Scott, 2012). The resulting phrases were then investigated, and the most frequent erroneous phrases were selected for further analysis.

### [5.2.1] Intra-L1-group similarities

To illustrate intra-L1-group similarities, Jarvis (2000) mentions Selinker (1992), according to whom a group of Hebrew-speaking learners of English tend to place adverbs before the object in English sentences (e.g. I like very much movies). The patterns of the (mis) use of English prepositions by Czech EFL learners were pointed out in this paper's section 5.1. Nevertheless, all errors in the use of independent prepositions, which appeared to be the most frequent type of preposition error in the speech of Czech EFL learners, were examined in greater detail. Using WordSmith Tools 8.0 (Scott, 2012) once more, errors in the use of independent prepositions were investigated, and their context was inspected to see which phrases are the most difficult for Czech EFL learners – see Figure 1 below.

**Figure 1** Errors in the use of independent prepositions across proficiency levels.



**Note:** The first category includes all days of the week.

The graph reveals that Czech EFL learners across all proficiency levels frequently err in the phrases *at school*, *on Monday*, *in the picture* and *at university*. These phrases will be examined more closely in the following part of the paper because, according to Jarvis (2000), intra-L1-group homogeneity is most evident when directly compared with inter-L1-group heterogeneity. Therefore, four other corpora of speech produced by EFL learners with diverse mother tongues were investigated.

### [5.2.2] Inter-L1-group differences

In order to explore whether specific types of preposition errors are caused by L1 transfer and are therefore L1 group-related or whether EFL learners in general make errors of these types, the Louvain International Database of Spoken English Interlanguage (LINDSEI) was examined. LINDSEI is a corpus of informal interviews with EFL learners of English from 11 mother tongue backgrounds (for more details, see <https://uclouvain.be/en/research-institutes/ilc/cecl/lindsei-cd-rom-and-handbook.html>).

Four LINDSEI subcorpora were selected for further investigation. Even though proficiency was initially defined with the institutional approach (most LINDSEI subjects are university undergraduates studying English), a random sample of five interview extracts from each of the eleven subcorpora was submitted to a professional rater, who was asked to rate them on the basis of the Common European Framework of Reference for Languages (CEFR) descriptors for speaking (Council of Europe, 2001). The results show that while some LINDSEI subcorpora clearly qualify as advanced, having a majority of C1 or C2 scores (cf. the Dutch, German, and Swedish subcorpora), others are rather in the higher intermediate range (Gilquin et. al., 2010). French, Italian, Japanese, and Spanish subcorpora might be labeled as B2 rather than advanced. For this reason, these specific subcorpora were selected for further examination because they correspond most closely to the analyzed data in the sample; the L1s were carefully selected in an attempt to control the proficiency variable, which may affect (as shown by the current analysis) learner productions. Moreover, the carefully selected comparable corpora comprise 50 informal interviews with EFL learners and a total of 34,116 words (Japanese subcorpus) to 83,294 words (French subcorpus), while my corpus comprises 45 interviews and totals 37,650 words (B turns only). For the above-mentioned reasons, the four selected corpora might be considered comparable to the corpus compiled for the purpose of the current research.

Only the most frequent erroneous phrases extracted from the Czech corpus under investigation with the help of WordSmith Tools 8.0 were further scrutinized in the other corpora to see whether specific types of preposition errors might be caused by L1 transfer and therefore be L1 group-related or whether EFL learners in general make errors of these types. The instances of erroneous prepositions were further broken down into the following groups: (1) school, (2) university, (3) picture, (4) days of the week, and (6) abroad.

The five corpora under investigation were examined using WordSmith Tools 8.0, with a focus on the most frequent words collocating with the incorrect prepositions that Czech EFL learners use. The results were then tested to verify whether the differences between the analyzed corpora were statistically significant. There is an appendix at the

end of this article showing the results of the statistical tests conducted at this stage of this research (Table 1 – Table 10).

## School

**Table 5** In/correct uses of the phrase at school across mother tongues.

	Correct	Incorrect
FR	<B> ...I had people in my classes . <b>at (er) secondary school</b> . who: who could have . (eh) become . (er) sportive sport-ives ...</B> FR022	<B>... who was a teacher <b>in high school</b> in Switzerland somewhere . in in (eh) not so far away from Zurich </B> FR021
IT	<B> ... I went with my with the school <b>at the high school</b> I went to: (er) Germany Austria and: (eh) ...</B> IT011	<B> ...I (eh) I studied (mm) <b>in (mm) in a school (erm)</b> . of of music but (mm) ... </B> IT012
JP	<B> he is a teacher <b>at junior high &lt;overlap /&gt; school</b> </B> JP010	<B> (erm) . I have one . brother and he is second grade <b>in high school</b> . and I have parents ... </B> JP006
SP	<B> ... they are finished the studies <b>at high school</b> and so (erm) they: they meet (erm) (er:) Maribel Verdú (er) who co= who comes from Spain .... </B> SP039	<B> I: <b>in the school</b> the classes were very boring </B> SP032
CZ	<A> how often do you speak English </A> <B> Not that often <b>only at school</b> </B> CZTT11	<B> .... so that's why I'm here because <b>on the secondary school</b> I was studying also information technologies ... </B> CZAI11

**Note:** The original coding system was used to code the correct and erroneous phrases (FR = French, IT = Italian, SP = Spanish, JP = Japanese). CZ... = Czech.

As evidenced by the above table, not only Czech EFL learners use the phrase *at school* incorrectly. The results of Pearson's chi-square test of independence show that the differences between the above-mentioned corpora are statistically significant ( $p = 0.001$ ). Japanese EFL learners make significantly more errors than Czech, Italian, and French learners in the phrase *at school*. Moreover, Czech EFL learners err more often than Italian and French learners in this specific phrase.

Thanks to the high frequency of occurrence of the word *school*, differences between the use of the prepositions *in*, *on* and *at* could be analyzed in greater detail. The results indicate that Japanese EFL learners use the preposition *in* more frequently than Italian, French, or Czech learners while Czech learners use the incorrect preposition *on* in this specific phrase significantly more often than any other nationality.

## University

All examined learner populations seem to struggle with the preposition preceding the word “university”. In examples (8) to (12), the majority of EFL learners substitute the correct preposition *at* with the incorrect preposition *in*, while a Japanese EFL learner (example 12) fails to use the preposition where it is obligatory.

(8) <A> how often do you study </A>

<B> like two three times per a week **on university** it's like really I just have to because it's just on me if I wanted to study I have to study ... </B> CZFIN9

(9) <B> (er) . I was already: in (er) . **in university** I . in fact I wanted to become a translator in English ... </B> FR016

(10) <B> ... I liked it very much . (er) now I'm enrolled **in university** this is my fifth year in university and last year I went to Spain because ... </B> IT045

(11) <B> in Manchester **in University of Salford** </B> SP045

(12) <B>... my father works **university** and my sister ... </B> JP013

However, the results of Pearson's chi-square test (see Tables 4 and 5 in the Appendix at the end of this article) show that the corpora subjected to scrutiny do not indicate statistically significant differences ( $p = 0.387$ ) in the learners' incorrect use of the phrase *at (the) university*.

### Picture

The current analysis showed that Czech EFL learners err in the use of the preposition preceding the word picture, and a closer analysis of the remaining sub-corpora revealed that French, Japanese, and Spanish learners also err in this particular phrase – see the table below.

**Table 6** Examples of some correct and incorrect uses of English prepositions in the analyzed sub-corpora.

	Correct	Incorrect
FR	<B> .... I think . yes well . not <b>in this picture</b> she really looks like the portrait but <laughs> </B> FR040	<B> and (er) well he starts drawing her . and then (er) <b>on the second picture</b> we see that she .... </B> FR015
IT	<B> ... and in the fourth . <b>in the fourth picture</b> she is at home I think and (eh) she shows... IT004	
JP	<B> there are (eh) . two people <b>in the &lt;overlap /&gt; picture</b> . and a man . is (em) . is drawing ... </B> JP001	<B> ...(er) beautiful . (er) . her hair style <b>on the picture</b> . is . (er) curling ... lastly (eh) the model... </B> JP023
SP	<B> much prettier than . <b>in the second . picture</b> in the first picture sorry . yes </B> SP001	<B> ... well okay the . snobbish woman <b>on the first picture</b> . wanted a famous . painter to: . make a picture ... SP005
CZ	<B> Well in <b>in the picture</b> is girl lying on her bed in the background there are three posters one of cat and two of probably her favourite singers </B> CZAI12	<B> ... at two of them are some men in the <b>on the picture</b> in the middle you get guitar and on the right singing and <b>on the picture</b> on the left there's nice cat nice white cat... </B> CZAI09

**Note:** The original coding system was used to code the correct and erroneous phrases (FR = French, IT = Italian, SP = Spanish, JP = Japanese). CZ... = Czech.



Moreover, tests show that the differences between the corpora under investigation in the use and accuracy of the phrase *in the picture* are statistically significant ( $p = 0.000$ ). French learners err more frequently than Italian, Japanese, Spanish, and Czech EFL learners in this specific phrase.

### **Days of the week**

All seven days of the week were investigated separately using WordSmith Tools 8.0. However, the individual frequencies (Monday, Tuesday, etc.) were grouped together into the ‘days of the week’ category and examined as a whole for the purpose of the statistical tests.

Some corpora had to be omitted from the testing of the days of the week and their prepositions due to the low occurrence of these phrases. For this reason, only Czech, French and Spanish corpora were analyzed in greater detail and tested using Pearson’s chi-square test of independence – see Table 9 in the Appendix of this paper.

The results indicate that the difference between these sub-corpora can be considered statistically significant ( $p = 0.000$ ) and that Czech EFL learners make more errors than Spanish or French learners in the use of the preposition *on* before the days of the week.

### **Abroad**

Czech EFL learners seem to be the only learner population using the word *abroad* incorrectly. It is therefore the only population with a frequency other than zero in the inaccurate use of this word. For this reason, the Czech corpus was compared to individual corpora using Fisher’s exact test, which revealed that the p-value was lower than the selected level of significance for all corpora except French, i.e. that Czech EFL learners make statistically significantly more errors in the use of *abroad* than any other learner population.

### **[5.2.3] L1-IL performance similarities**

An obvious motivation for selecting an incorrect preposition is the learner’s L1, i.e. Czech EFL learners may have chosen a certain preposition because its sense or form corresponds to that of a Czech preposition used in the same context. For example, a plausible explanation for the preposition choice in ‘on the picture’ is that the Czech language requires the preposition ‘na’ in a similar context (na škole). English ‘on’ and Czech ‘na’ are basic correspondents in many contexts (on the table → na stole, on a farm → na farmě) in the sense that they share the same basic meaning.

According to Nesselhauf (2003, p. 234), similarity is considered an indication that influence is likely. Therefore, the deciding factor indicating a possible negative transfer adopted in the current research is the degree of linguistic correspondence between English and Czech. To uncover the possible indications of L1 transfer, the most frequent inaccurate prepositions were subsequently inspected in terms of their ‘match’ between the L1 and TL, i.e. the syntactic structures required by the two languages in the specific context were investigated, as well as the correspondence of the basic meanings. To show whether Czech-speaking learners of English parallel their use of prepositions in their L1,

the same phrases that learners produce incorrectly were analyzed in the learners' mother tongue. The results are hoped to reveal what in the L1 motivates the IL behavior.

**Table 7** Erroneous and correct phrases and their direct translation equivalents in Czech.

Erroneous phrases	Correct phrases
on school → na škole	at school
in abroad → v zahraničí	abroad
in Monday → v pondělí	on Monday
on the picture → na obrázku	in the picture
on the university → na univerzitě	at the university

Even though the above analysis showed that not only Czech EFL learners, but also EFL learners in general make the majority of the above-mentioned specific errors in the use of English prepositions, L1-IL equivalence in form was found for all the erroneous phrases, while a direct translation equivalent of the Czech phrases does not seem to correspond to the majority of the correct English phrases. For this reason, the L1 may prompt learners to use prepositions that are erroneous and very atypical in English, such as *in Monday*.

Some of the above-mentioned erroneous phrases (in abroad, on school) only appeared in the Czech corpus, where the correct preposition is almost always replaced by an incorrect one, corresponding more to the Czech equivalent. And since the learners' use of some L2 features can be shown to parallel their use of a corresponding L1 feature, it can be hypothesized quite reasonably that the L1 influences a Czech learner's IL. On the positive side, while L1 influence may be negative in cases of non-correspondence, it may equally be positive in cases where the prepositions in L1 and L2 correspond. However, the investigation of positive transfer was not intended to form a part of the present research, even though such an investigation might prove helpful in future research.

Nevertheless, it should be noted that it is impossible to remove all ambiguity concerning whether negative transfer took place, as a complete investigation of negative transfer would require extra data, which the current corpus alone does not provide.

## [6] Conclusions

The current study adds to existing knowledge of preposition accuracy (Chon et al., 2021; Granger, 2003; Gráf, 2015; Thewissen, 2015) by affirming the prevailing view of prepositions as a formidable challenge for L2 learners. The findings show that accuracy in the use of English prepositions increases with learners' increasing proficiency level, supporting the position that negative L1 transfer plays an important role in the production of EFL learners.

Overall, the findings from this study help us to better understand the construct of prepositional accuracy in naturalistic speech produced by EFL learners. Unlike past studies, this study focuses specifically on the accuracy order of English prepositions and supports the notion that prepositional accuracy on the part of speakers might be a good

predictor of language proficiency. The present research has revealed a clear hierarchy of preposition errors regarding their order of frequency across proficiency levels for the Czech EFL speech production scrutinized. Considering this, it seems reasonable to conclude that preposition errors might be the most frequent among A1 learners of EFL and the least frequent among C2 learners, which makes them good predictors of the accuracy of L2 speech production. However, future research should concentrate on determining whether the hierarchy of preposition errors across proficiency levels found in this paper remains stable for subjects with lower (A1) and higher (C1 and C2) levels of proficiency in the target language. It would be interesting to know if and how the order of frequency of preposition error types changes across all six proficiency levels.

Contrarily, preposition errors do not seem to be strong predictors of EFL speech quality if this quality is measured in terms of communicative effectiveness (the majority of preposition errors include prepositions *in*, *on* and *at*, which hardly prevent understanding in communication). Since the majority of preposition errors did not have the potential to hinder understanding, their inaccurate use should not be considered a serious problem. Nevertheless, the only goal of communication should not be to get the message across – there exists a danger of ‘pidginizing’ the English language (Mukherjee & Rohrbach, 2006, p. 210).

Considering that the presence of preposition errors does not play any relevant role in the quality of L2 speech, we can conclude that the correlation between accuracy and speech quality does not hold when only preposition errors are considered. Given the loose link between preposition accuracy and speech quality, it seems reasonable to encourage learners to use prepositions, as this is likely to have positive effects, even if there are some errors in the use of them. Current theories (e.g. Information Processing – see, for example, Ellis and Robinson, 2008) emphasize the role of exposure and meaningful practice as factors that will move learners forward and allow them to get rid of mistakes, even if backsliding can commonly occur. From the pedagogical point of view, the development of learner strategies, especially memory-related and cognitive ones, can be recommended (see, for example, Norton & Toohey, 2001; Taylor, 1975; or Thompson, 2005).

Furthermore, the data scrutinized for the current study support the usefulness of comparisons with Czech as L1, which can therefore be recommended on the basis of the findings.

All in all, my empirical data affirm that even though both the Czech language and the English language are Indo-European languages with similar structures for preposition use and a large number of formally similar preposition lexemes, it might be the basic correspondents in different contexts for preposition use that cause trouble. Following this, the high degree of inaccurate use of English prepositions may be viewed as a result of negative influence, although it is impossible to verify this with a high degree of certainty based on the present material. An area for future investigation, therefore, concerns reasons for inaccurate preposition use that may/may not plausibly be attributed to negative transfer. For an example of a study on negative transfer see Kapranov (2020), who investigated intermediate EFL students’ self-assessment of phonetically difficult words in

English. In his study, the participants were asked to comment on what they considered to be a possible cause of the difficulty, and the results of his study indicate that phonetically difficult words are associated with the segmental elements that are absent in the phonological inventory of the participants' first language.

A limitation of my study is that only one mother tongue background came under scrutiny in its totality. Furthermore, as pointed out above, an unfortunate side effect of not being able to customize the data according to proficiency levels was that I obtained a different amount of data in the groups at the six levels of proficiency, which mainly precluded valid comparisons from the proficiency perspective, including the marginal levels of proficiency (A1 and C1–C2). Nevertheless, Gráf (2015, p. 116), who examined accuracy in the speech of Czech advanced learners of English, focused – among many other things – on dependent and independent prepositions in his analysis. Of the total number of 121 errors in the use of independent prepositions in the corpus he examined, 105 errors (87%) involved the prepositions *in*, *at* or *on*. These instances were further broken down into the following groups: (1) on the picture / painting / drawing / portrait; (2) in university / school; (3) various other instances. The results of the current analysis (which focused on A2, B1, and B2 learners) seem to correspond to the results of Gráf's (2015) analysis (which focused on advanced learners) and indicate that learners across all proficiency levels, including advanced learners, make similar errors in the use of English prepositions, albeit less frequently.

On the positive side, regardless of the potential limitations of the current study, it seems to provide important assessments of preposition accuracy that are not available through other means. These assessments provide us with a fuller picture of the elements that help define preposition accuracy in speech produced by EFL learners. Moreover, the results of my research suggest interesting areas for future investigation, which would certainly yield valuable findings in addition to those already posted. For example, although repetition was not intended to form a part of the present research, its examination might prove helpful in future research since the frequency of repetition may affect the interpretation of results, mainly in relation to the overuse of certain prepositions in the speech of EFL learners. Additionally, further investigation of sources of errors might prove helpful in future research. A way of determining a possible error source might be to consult the learner, i.e. to identify what processes learners invoke when they do not know the TL form. If the learner can identify the source of their specific error, these can be reported and added to the existing knowledge of the learner's language.

Overall, research on the preposition errors EFL learners make is a rich area of study aiming to understand the nature of these errors, their underlying causes, and effective ways to address them in EFL teaching and learning contexts.

Finally, advances in the field of LCR deserve special praise, particularly the WordSmith Tools program (Scott, 2012), which facilitated observations and enabled me to search for linguistic patterns. To this end, WordSmith made observing the data qualitatively a less tedious task. CLAWS4 and oTranscribe, which appeared to be highly consistent and accurate, also facilitated work on this article.

## [Notes]

- 1 According to Rayson et al. (2001, p. 303), the use of prepositions differs between speech and writing more than is the case with most other word classes, and a high proportion of preposition use is associated with the informative and nominal tendency of written language  
Bentley, E. (2013). oTranscribe: A free web app to take the pain out of transcribing interviews. Retrieved from <https://otranscribe.com/>  
Heywood, J., & Pouliquen, B. (2003). CLAWS4: The tagging of the British National Corpus. In Proceedings of the corpus linguistics 2003 conference (pp. 171–180).  
Scott, M., & Baker, P. (2011). WordSmith Tools version 8.0 [Computer software]. Retrieved from <https://www.lexically.net/wordsmith/version8/>

## [Bibliography]

- Bachman, L. F., and Palmer, A. S. (1996). *Language Testing in Practice: Designing and Developing Useful Language Tests*. Oxford and New York: Oxford University Press.
- Bardovi-Harlig & Sprouse. (2018). Negative Versus Positive Transfer. In J. I. Lontas (Ed.) *The TESOL Encyclopedia of English Language Teaching*.
- Biber, D., Conrad, S., and Leech, G. (2002). *Longman student grammar of spoken and written English*. Harlow: Pearson.
- Biber, D., Johansson, S., Leech, G., Conrad, S., and Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson.
- Brand, C., and Götz, S. (2011). Fluency versus Accuracy in Advanced Spoken Learner Language: A Multi-Method Approach. *International Journal of Corpus Linguistics*, 16, 255–275.
- Carter, R., and McCarthy, M. (2006). *Cambridge grammar of English*. Cambridge: Cambridge University Press.
- Celce-Murcia, M., and Larsen-Freeman, D. (1999). *The Grammar Book: An ESL/EFL Teacher's Course (2nd ed.)*. Boston, MA: Heinle and Heinle.
- Chodorow, M., Tetreault, J., and Han, N.-R. 2007. 'Detection of grammatical errors involving prepositions', 4th ACL-SIGSEM Workshop on prepositions, Prague, Czech Republic.
- Chon, Y., Shin, D., and Kim, G. (2021). Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System*, 96. Retrieved from: <https://doi.org/10.1016/j.system.2020.102408>.
- Cogo, A. and Dewey, M. (2012). *Analysing English as a Lingua Franca: A Corpus-driven Investigation*. London, New York: Continuum International Publishing Group.
- Corder, S.P. (1981). *Error Analysis and Interlanguage*. Oxford: Oxford University Press.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.
- De Felice, R., and Pulman, S. (2009). Automatic Detection of Preposition Errors in Learner Writing. *CALICO Journal*, 26(3), 512–528.

- Dose-Heidelmayer, and S., Götz, S. (2016). The progressive in spoken learner language: A corpus-based analysis of use and misuse. *International Review of Applied Linguistics in Language Teaching*, 54(3), 229–256.
- Ellis, R. , and Barkhuizen, G. (2005). *Analysing learner language*. Oxford and New York: Oxford University Press.
- Fleiss, J. L., Levin, B., and Paik, M. C. (2013). *Statistical methods for rates and proportions*. New York, NY: John Wiley & Sons.
- Gass, S. M., and Selinker, L. (2008). *Second Language Acquisition: An Introductory Course*. New York and London: Routledge.
- Gilquin, G., De Cock, S., and Granger, S. (2010). *Louvain International Database of Spoken English Interlanguage*. Presses universitaires de Louvain: Louvain-la-Neuve.
- Gráf, T. (2015). *Accuracy and fluency in the speech of the advanced learner of English*. Praha: Univerzita Karlova v Praze.
- Granger, S. (2003). Error-tagged learner corpora and CALL: A promising synergy. *CALICO Journal*, 20 (3), 465–480.
- Granger, S., Swallow, H., and Thewissen, J. (2022). *The Louvain Error Tagging Manual. Version 2.0*. CECL Papers 4. Centre for English Corpus Linguistics: Louvain-la-Neuve.
- Greenbaum, S., and Quirk, R. (1990). *A Student's Grammar of the English Language*. London: Longman.
- Gries, S. Th., and M. Paquot. (2020). Writing up a corpus-linguistic paper. In M. Paquot & S. Th. Gries (eds.). *Practical Handbook of Corpus Linguistics*.
- Jarvis, S. (2000). Methodological Rigor in the Study of Transfer: Identifying L1 Influence in them Interlanguage Lexicon. *Language Learning*, 50(2), 245–309.
- Kapranov, O. (2020). Intermediate EFL students' self-assessment of phonetically difficult words in English. *Ostrava Journal of English Philology*, 12(2), 69–91.
- Mauranen, A. 2012. *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge: Cambridge University Press.
- Mukherjee, J., Rohrbach, J., Kettemann, B., and Marko, G. (2006). *Rethinking applied corpus linguistics from a language-pedagogical perspective: New departures in learner corpus research*. Frankfurt: Peter Lang. Cognitive Linguistics Bibliography.
- Nacey, S., and Graedler, A.L. (2015). Preposition use in oral and written learner language. *Bergen Language and Linguistic Studies*, 6, 45–62.
- Nesselhauf, N. (2003). The Use of Collocations by Advanced Learners of English and Some Implications for Teaching. *Applied Linguistics*, 24(2), 223–242.
- Norton, B., and Toohey, K. (2001). Changing perspectives on good language learners. *TESOL Quarterly*, 35(2), 307–322.
- Parrott, M. (2000). *Grammar for English Language Teachers*. Cambridge University Press.
- Quirk et al. (1985). *A Comprehensive Grammar of the English Language*. London: Pearson Longman.
- Rayson, P., Wilson, A., and Leech, G. (2001). Grammatical word class variation within the British National Corpus Sampler. *Language and Computers*, 36(1), 295–306.

- Robinson, P., and Ellis, N. C. (2008). *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York and London: Routledge.
- Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics*, 24, 209–239.
- Taylor, B. P. (1975). Adult language learning strategies and their pedagogic implications. *TESOL Quarterly*, 9, 391–399.
- Thewissen, J. (2015). *Accuracy across proficiency levels: A Learner corpus approach*. Belgium: Presses universitaires de Louvain.
- Thompson, S. (2005). *The Good Language Learner*. Birmingham University. Retrieved from:  
<https://www.birmingham.ac.uk/documents/college-artslaw/cels/essays/secondlanguage/essaygllsthompson.pdf>
- Thewissen, J. (2013). Capturing L2 Accuracy Developmental Patterns: Insights From an Error-Tagged EFL Learner Corpus. *The Modern Language Journal*, 97, Supplement. Capturing Second Language Development Through Learner Corpus Analysis, 77–101.

**[Address]**

Masaryk University  
Gorkého 57/7  
60200 Brno  
[zdenka.neumanova@email.cz](mailto:zdenka.neumanova@email.cz)

*Zdeňka Neumanová is a doctoral student in Experimental and Applied Linguistics at the Department of English and American Studies, Masaryk University, Brno. In her dissertation, which focuses on an analysis of spoken learner production in English, she deals with grammatical errors in the speech of Czech university learners of English. Her wider research interests include (but are not limited to) corpus linguistics, English as a Foreign Language (EFL), error analysis, Foreign Language Teaching (FLT), learner corpus research, Second Language Acquisition (SLA), and Transfer/Cross-Linguistic Influence (CLI).*

**[Appendix]****Table 1** In / on / at school corpora accuracy cross tabulation.

<b>Corpus * accuracy cross tabulation</b>						
			Accuracy		Total	
			Incorrectly	Correctly		
Corpus	Czech	Count	38	22	60	
		% within corpus	63.3%	36.7%	100.0%	
	French	Count	21	24	45	
		% within corpus	46.7%	53.3%	100.0%	
	Italian	Count	8	17	25	
		% within corpus	32.0%	68.0%	100.0%	
	Japanese	Count	14	3	17	
		% within corpus	82.4%	17.6%	100.0%	
	Spanish	Count	29	10	39	
		% within corpus	74.4%	25.6%	100.0%	
	Total		Count	110	76	186
			% within corpus	59.1%	40.9%	100.0%

**Table 2** In/on/at school Pearson's chi-square test results.

<b>Chi-Square Tests</b>			
	Value	Df	Asymptotic Significance (2-sided)
Pearson Chi-Square	18.483 <sup>a</sup>	4	.001
N of Valid Cases	186		

**Table 3** Pearson chi-square test results for the differences between the use of specific prepositions.

<b>Chi-Square Tests</b>			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	45.784 <sup>a</sup>	8	.000
N of Valid Cases	186		



**Table 4** In / on / at the university Pearson chi-square test results.

<b>Corpus * accuracy cross tabulation</b>					
			Accuracy		Total
			Incorrectly	Correctly	
Corpus	Czech	Count	6	3	9
		% within corpus	66.7%	33.3%	100.0%
	French	Count	4	8	12
		% within corpus	33.3%	66.7%	100.0%
	Italian	Count	8	16	24
		% within corpus	33.3%	66.7%	100.0%
	Japanese	Count	1	1	2
		% within corpus	50.0%	50.0%	100.0%
Spanish	Count	10	9	19	
	% within corpus	52.6%	47.4%	100.0%	
Total		Count	29	37	66
		% within corpus	43.9%	56.1%	100.0%

**Table 5** Pearson chi-square test results for the phrase at the university.

<b>Chi-Square Tests</b>			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	4.144 <sup>a</sup>	4	.387
N of Valid Cases	66		

**Table 6** Accuracy of the phrase in the picture across the analyzed corpora.

<b>Corpus * accuracy cross tabulation</b>						
			Accuracy		Total	
			Incorrectly	Correctly		
Corpus	Czech	Count	8	52	60	
		% within corpus	13.3%	86.7%	100.0%	
	French	Count	6	2	8	
		% within corpus	75.0%	25.0%	100.0%	
	Italian	Count	0	84	84	
		% within corpus	0.0%	100.0%	100.0%	
	Japanese	Count	2	11	13	
		% within corpus	15.4%	84.6%	100.0%	
	Spanish	Count	13	99	112	
		% within corpus	11.6%	88.4%	100.0%	
	Total		Count	29	248	277
			% within corpus	10.5%	89.5%	100.0%

**Table 7** Pearson chi-square test results.

<b>Chi-Square Tests</b>			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	46.379 <sup>a</sup>	4	.000
N of Valid Cases	277		

**Table 8** Accuracy of the preposition *on* preceding days of the week.

<b>Corpus * accuracy cross tabulation</b>						
			Accuracy		Total	
			Incorrectly	Correctly		
Corpus	Czech	Count	12	7	19	
		% within corpus	63.2%	36.8%	100.0%	
	French	Count	0	33	33	
		% within corpus	0.0%	100.0%	100.0%	
	Italian	Count	0	4	4	
		% within corpus	0.0%	100.0%	100.0%	
	Japanese	Count	0	1	1	
		% within corpus	0.0%	100.0%	100.0%	
	Spanish	Count	0	7	7	
		% within corpus	0.0%	100.0%	100.0%	
	Total		Count	12	52	64
			% within corpus	18.8%	81.3%	100.0%

**Table 9** Pearson chi-square test results.

<b>Chi-Square Tests</b>			
	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	31.713 <sup>a</sup>	2	.000
Likelihood Ratio	34.590	2	.000
Linear-by-Linear Association	11.399	1	.001
N of Valid Cases	59		

**Table 10** Accuracy of the word *abroad*.

<b>Corpus * accuracy cross tabulation</b>						
			Accuracy		Total	
			Incorrectly	Correctly		
Corpus	Czech	Count	6	5	11	
		% within corpus	54.5%	45.5%	100.0%	
	French	Count	0	4	4	
		% within corpus	0.0%	100.0%	100.0%	
	Italian	Count	0	17	17	
		% within corpus	0.0%	100.0%	100.0%	
	Japanese	Count	0	7	7	
		% within corpus	0.0%	100.0%	100.0%	
	Spanish	Count	0	23	23	
		% within corpus	0.0%	100.0%	100.0%	
	Total		Count	6	56	62
			% within corpus	9.7%	90.3%	100.0%