Univerzita Karlova

Filozofická fakulta

Ústav anglického jazyka a didaktiky

Filologie – Anglický jazyk

DISERTAČNÍ PRÁCE

Mgr. Lucie Jiránková

The development of L2 inflectional morphology and cross-language interference effects

Vývoj flektivní morfologie druhého jazyka a mezijazykové interferenční efekty

Vedoucí práce: Luca Cilibrasi, Ph.D.

2021

ACKNOWLEDGEMENTS

I would like to use this opportunity to thank several people without whose help and support you would not be reading this dissertation right now. Being in the last stage of my Ph.D. and reflecting back on these past four years, I know those have not only been the most challenging years of my life but also the richest academically. And I would not be able to gain such experience without my colleagues, friends, and family.

First and foremost, I owe a huge thank-you to my supervisor, Luca Cilibrasi, Ph.D., for his constant guidance, helpful comments, and encouraging approach. I meant it when I said that he was the best supervisor a Ph.D. student could wish for. Working with him has been one of the highlights of my academic career so far, as he is not only a great academic and a true professional but also an empathic person.

I am also very thankful to PhDr. Tomáš Gráf, Ph.D., who always found time to consult any pedagogical implications of my thesis and offered invaluable insights into the practical application of my data. His pedagogical guidance will always be very dear to me since he cared about me and supported me during the entire course of my studies.

I also highly value any constructive feedback offered by my departmental colleagues, who were always willing to help and offer assistance. Special appreciation goes to PhDr. Pavlína Šaldová, Ph.D., for submitting my proposal to the ethical committee and securing funding for the experiments.

There were several other researchers who significantly contributed to this work and provided helpful insights into any methodological or statistical struggles. Doc. Mgr. Radek Skarnitzl, Ph.D., offered his assistance with Praat and booked the recording studio. Mgr. Pavel Šturm, Ph.D., provided me with his recent paper on the phonotactic analysis of Czech words. Dr. Ryan Blything, Prof. Adam Albright, and Prof. Bruce Hayes gave me permission to use their novel word stimuli for two of my experiments. Suzanne Lewis, M.A. then recorded the stimuli in the recording booth. Dr. Ana I. Pérez Muñoz offered valuable insights into my statistical queries. Many Faculty of Arts employees assisted with the recruitment of participants and with booking the rooms used for the experiments, and I will be forever grateful to them.

This project has also been financially supported by a Specific Academic Research grant, "Jazyk a nástroje pro jeho zkoumání."

However, completing this dissertation required more than the support from my academic colleagues. My indescribable appreciation goes to my family, especially to my mum, who was showering me with unconditional love and support for the past four years. I have no idea how she put up with my crankiness, sleep deprivation, and coffee highs, but she somehow did, and I could not be more grateful. A huge thank-you also goes to my wine club that provided me with immeasurable support and help and also to my best friend, who was the best support system a Ph.D. student could wish for and who regularly kept my sanity intact by distracting me with Asian TV shows.

Last but not least, I am also incredibly indebted to all the participants. Their willingness to participate stood at the core of this dissertation, and without their support, it would not have been possible to write this thesis.

DECLARATION

Prohlašuji, že jsem tuto disertační práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného či stejného titulu.

I declare that the following dissertation is my own work for which I used only the sources and literature mentioned and that this thesis has not been used in the course of other university studies or in order to acquire the same or another type of diploma.[1]

Jiránková

V Praze dne 10. července 2021 ……………………………….

Mgr. Lucie Jiránková, v. r.

[1] Parts of this thesis have been previously published in peer-reviewed journals. For more information, see the footnotes at the beginning of each experimental chaper.

PERMISSION

Souhlasím se zapůjčením této disertační práce ke studijním účelům.

I have no objections to this dissertation being borrowed and used for study purposes.

ABSTRACT

The present dissertation investigates the **development** in the production and perception of **inflectional morphology in second-language learners of English** and the role of their mother tongue during this development. The data analysed in this thesis stem from three psycholinguistic experiments that examine the production and comprehension of English novel words (thus investigating the sublexicon without the activation of word meaning). The **first experiment** focuses on the perception of inflectional morphemes in English novel words in L2 students at the A0 to C1 proficiency levels. Reaction-times analysis has shown that L2 learners seem to be (similarly to native speakers (e.g., Post et al., 2008)) sensitive to the presence of morphosyntactic information at the sublexical level, and they appear to decompose inflected forms into stems and affixes during perception and conduct an implicit phonetic analysis of the stem. The presence of these patterns across all levels suggests that L2 performance might be influenced by L1: Czech is morphologically much richer than English, and Czech speakers might thus be in general sensitive to morphological analysis of words. The **second experiment** investigates the production of inflected forms, more specifically those of past tense, in L2 learners of English at the A1 to C1 levels (and in a control group of English native speakers) with an elicitation task. The analysis of produced forms has shown that the L2 learners display a progressive development from the application of default rules (at the A1 to B1 levels) to the use of analogy (in more proficient B2 and C1 levels), a procedure that is used productively by native speakers. Reaction-times analysis has shown that L1 Czech functions as a facilitator at the lowest language levels (items that are phototactically legal in both Czech and English are analysed more quickly). Given the findings of the first two experiments (pointing in partly different directions), I suggest the presence of redundant processing, which enables parallel activation of two (redundant) mechanisms. The **third experiment** focuses on language transfer, using a lexical decision task with the novel verbs from Experiment 2, paired with existing English words, and it assesses L2 learners distributed on a proficiency scale. Reaction-times analysis has shown that Czech has a mild facilitatory effect on performance, not only for the lowest language levels but also for more proficient learners, suggesting that transfer effects are not limited to non-proficient speakers. Altogether, the findings of this thesis offer complex look into how inflectional morphology develops in L2 learners from the lowest to the proficient levels and also into the role of L1. In the last chapter, I additionally discuss possible **pedagogical implementations** of these findings into an L2 classroom.

Key words: inflectional morphology, interference effects, production, perception, language transfer, second-language acquisition, rules, analogy, morpheme storage, lexicon organization, Czech, English

ABSTRAKT

Předkládaná disertační práce zkoumá vývoj produkce a percepce flektivní morfologie u studentů angličtiny jako druhého jazyka a roli mateřského jazyka na tento vývoj. Data analyzovaná v této práci pocházejí ze tří psycholingvistických experimentů, které zkoumají produkci a porozumění anglickým neslovům (zkoumají tedy sublexikon, bez aktivace slovního významu). **První experiment** se zaměřuje na vnímání flektivních morfémů v anglických neslovech u studentů angličtiny na úrovních A0 až C1. Analýza reakčních časů ukázala, že studenti angličtiny se zdají být (podobně jako rodilí mluvčí (např. Post et al., 2008)) citliví na přítomnost morfosyntaktických informací na sublexické úrovni, a pravděpodobně během percepce rozkládají flektivní formy slov na kmeny a přípony a zároveň provádějí implicitní fonetickou analýzu kmene. Podobné reakce na všech úrovních naznačují, že studenti angličtiny mohou být ovlivněni svým mateřským jazykem: čeština je morfologicky mnohem bohatší než angličtina a čeští mluvčí tak mohou být obecně citlivější na morfologickou analýzu slov. **Druhý experiment** zkoumá produkci flektivních forem, konkrétně forem minulého času, u studentů angličtiny na úrovních A1 až C1 (a u kontrolní skupiny rodilých mluvčí angličtiny) pomocí elicitace. Analýza vyprodukovaných forem slov ukázala, že studenti angličtiny postupují od aplikace standardních pravidel (na úrovních A1 až B1) k využití analogie (na pokročilejších úrovních B2 a C1), kterou využívají i rodilé mluvčí. Analýza reakčních časů ukázala, že čeština funguje na nejnižších jazykových úrovních jako facilitátor (jelikož neslova, které byla fototakticky legální v češtině i angličtině, byla analyzována rychleji). Na základě zjištění prvních dvou experimentů (směřujících částečně odlišnými směry) navrhuji existenci redundantního zpracování jazyka, které umožňuje paralelní aktivaci dvou (redundantních) mechanismů. **Třetí experiment** se zaměřuje na jazykový přenos s využitím neslov z experimentu 2 a reálných anglických slov a s pomocí lexikálního rozhodování, do kterého byli zapojeni studenty angličtiny na škále jazykové pokročilosti. Analýza reakčních časů ukázala, že čeština má mírný facilitační účinek, a to nejen u nejnižší jazykové úrovně, ale i u pokročilejších studentů, což naznačuje, že vlivy jazykového transferu se neomezují pouze na nepokročilé mluvčí. Výsledky této práce nabízejí komplexní pohled na vývoj osvojování flektivní morfologie studentů angličtiny od nejnižší úrovně po pokročilé mluvčí a také na roli mateřského jazyka. Poslední kapitola nabízí možné **pedagogické využití** těchto poznatků v hodinách anglického jazyka.

Klíčová slova: flektivní morfologie, interferenční efekty, produkce, percepce, jazykový transfer, osvojování druhého jazyka, pravidla, analogie, ukládání morfémů, organizace lexikonu, čeština, angličtina

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANOVA	Analysis of variance
BF	Biphone frequency
BNC	British National Corpus
CELEX	Centre for lexical information
cond	Condition
CQL	Corpus query language
Df	Degrees of freedom
FQ	First quartile
GCM	Generalized Context Model
ID	Identification number
IQR	Interquartile range
IR	Irregular
L1	First (native) language
L2	Second language
MGL	Minimal Generalization Learner
MGL_IR	Similarity to existing English irregular verbs
MGL_R	Similarity to existing English regular verbs
morpho	Morphological
ms	Milliseconds
NS	Native speakers
non-morpho	Non-morphological
N / n	Number / sample size
part	Participant(s)
PP	Past participle
PSF	Positional segment frequency
R	Regular
RP	Received pronunciation
RT(s)	Reaction time(s)
RW	Real word
s	Seconds
SD	Standard deviation
SE	Standard error
Sq	Square
TQ	Third quartile
VIF	Variance inflation factor
VVB	Lexical verb: finite base Form

1. INTRODUCTION

Second-language morphology is the centre of attention of many theories and studies on second-language acquisition. Two areas of research have proved especially fruitful to its investigation: (1) the specificity of second-language morphology acquisition and the issues that are connected with it, and (2) its difference from the acquisition of morphology in the learner's mother tongue. So far, researchers have been unable to give a clear answer on which specific problems underlie second-language acquisition of morphology and how the mother tongue actually influences this acquisition. The mechanisms used in the processing of both monolingual and second-language morphology in real time and their effect on the organization of the mental lexicon have attracted strong research interest. Traditionally, two approaches are regarded as direct opponents: (i) the single-mechanism approach, positing that all morphologically inflected forms are processed through one single mechanism based either on analogy to stored forms (J. Bybee, 1995) or solely on the online use of morphosyntactic rules (Halle & Mohanan, 1985) and (ii) the dual mechanism (Prasada & Pinker, 1993; Chialant & Caramazza, 1995; Clahsen, 1999), positing that regular and irregular forms are processed differently – while irregular forms are retrieved as units from our memory, regular forms are produced online by default rules (Clahsen, 2006). Sometimes, a third "hybrid" mechanism of morphological processing is cited (e.g., Albright & Hayes, 2003), which "makes use of multiple, stochastic rules" (Albright & Hayes, 2003: 120), combining associative processes and phonological micro-rules into structural analogy.

An interesting approach to the study of morphology consists of investigating the processing of inflections in novel words, tapping into the sublexicon with no access to the lexical meaning. Only a few studies have focused on the study of L2 morphology at the sublexical level, employing only the generalization mechanism and focusing solely on the grammatical form (e.g., Murphy, 2000; Cuskley et al., 2015), and even fewer studies have attempted to describe both second-language morphological production and perception. In addition, previous studies usually focused on one proficiency group only, and very few have attempted to describe the development of second-language morphology from the initial stages of L2 learning to near proficiency (usually involving only two or three proficiency groups, e.g., Jiránková, 2017). This thesis, therefore, aims at (at least partially) filling this gap by investigating the development of inflectional morphology in second-language learners of

English with Czech as L1 (in comparison to native speakers of English) at different stages of proficiency, ranging from very beginners to nearly proficient speakers, and describe potential cross-language interference effects. Our main research area will be the storage of inflectional morphology and its online processing. In addition to that, we will also try to investigate the role L1 Czech might play in this acquisition. In order to do that, the thesis aims to focus both on perception and production of English inflectional morphology and attempts to bring new insights into a major and long-standing psycholinguistic debate about a speaker's ability to (a) perceive and (b) produce (novel) morphologically inflected forms (i.e., inflected novel words) and (c) the role of L1. To investigate the complexity of L2 morphology acquisition and to provide the most complex picture of its development, three experiments were carried out: (1) one concentrating on the sublexical perception of morphologically-inflected forms, (2) one concentrating on the production of novel morphologically-inflected verbs, and (3) one concentrating solely on possible cross-language interference effects from Czech.

The intended outcome of this thesis is twofold. Our findings will be relevant from the theoretical perspective since we should get a closer look at how the sublexicon operates for L2 learners of all language levels, how inflectional morphology is stored and processed by L2 learners of English, and how L1 Czech influences the performance. Implications for L2 acquisition of inflectional morphology will be addressed using the findings of our experiments. The results should also be relevant for pedagogical practice and L2 learning. Our findings might help us view L2 acquisition of inflectional morphology in a new light – this includes the specifics of L2 acquisition of inflectional morphology and the role of L1 Czech – and new methods of instruction may be proposed as a consequence.

This dissertation is, therefore, organized as follows: Chapter 2 provides the theoretical overview of inflectional morphology, its acquisition in L1 and L2, the specifics of acquiring L2 morphology, information on the two (alternatively three) models of processing hinted above, a brief overview of past research on L2 perception and production, along with the aims of the thesis, the means of their attainment, and hypotheses. The following three chapters then describe each experiment in detail. Given the substantial difference in methodologies, each experiment needed to be handled as a separate chapter. Chapter 3, therefore, introduces the perception experiment with a theoretical background specific to the experiment, our hypotheses, methodology, practical analysis, and the discussion of the findings. Chapter 3 illustrates how, similarly to native speakers of English, the perception of inflectional morphology in L2 learners shows signs of morphological decomposition while simultaneously showing some frequency effects. A similar structure is found in Chapter 4, which presents the production experiment, and Chapter 5, which provides a detailed description of the lexical decision task. Chapter 4 shows that the production of inflectional morphology in L2 learners is characterized by a progression from the application of rules to the use of analogy, which is used exclusively by native speakers. Chapter 5 shows that some minor Czech transfer effects might be at play not only for beginner L2 learners but also for more proficient participants. Chapter 6 then provides a section on a general discussion and conclusion, offering a brief overview of the findings, their relevance in association with previous studies, a reflection of the limits of the study, directions for future research, and a suggestion for a new teaching method concerning the acquisition of L2 past-tense morphology. In summary, this thesis aims to make theoretical contributions to the long-lasting psycholinguistic debates on the perception and production of L2 inflectional morphology and hopes to provide some ideas on how pedagogy can make use of them.

2. THEORETICAL BACKGROUND AND HYPOTHESES

2.1 Word-final positions and inflectional morphology

In order to outline its main features, inflectional morphology is often directly contrasted with derivational morphology (Crystal, 2011). According to Plag (2003), one of the most prominent differences between the two is that (English) inflectional morphology encodes "grammatical categories such as plural (*workers*), person (*works*), tense (*picked*), or case (*John's*)" (Plag, 2003: 19) and contributes to syntactic relevance (as opposed to derivational morphology). Another specific feature of (regular) inflectional morphology is its fixed word-final position for affixes, contrary to derivational morphology that presents affixes in different positions; only irregular forms use other means of inflection than suffixes, e.g., *man – men* or *drink – drank.* Inflectional morphemes also typically do not change the meaning or the part of speech of the base (e.g., *worker-s*), as is sometimes the case with derivational morphemes (e.g., *work-er*). Inflectional suffixes are also semantically transparent (i.e., the meaning of the whole inflected form is deduced through the semantic combination of its individual parts) and tend to be fully productive with fewer restrictions on their use with novel word forms (Plag, 2003).

Word-final positions are interesting both morphologically and phonologically. From the phonological point of view, in contrast to other word positions, these are prone to deletion and assimilation (Roach, 2009; Harris, 2011). For example, the final phoneme in *that* /ðæt/ may become assimilated to the first phoneme of the following word in rapid, casual speech, as in *that package* /ðæp pækɪdʒ/; or the middle plosive may altogether disappear in clusters of three plosives and *looked back* will become /lʊk bæk/ with the /t/ sound omitted (Roach, 2009). Pater (2007) and Harris (2011) show that the saliency of word-final positions, however, depends on whether any (potential) morphosyntactic information is present or not – if it is, even word-final position can be salient and avoid such processes as reduction or deletion (Harris, 2011).

In many languages, word-final positions are of high importance since they carry morphological information (Pater, 2004). Focusing solely on English, all of its bound inflectional morphemes appear at the end of words and consist of a relatively limited set of suffixes - *-(e)s, -(e)d, -ing, 's, -er, -est,* and *-en.* Inflectional suffixes *-(e)s* and *-(e)d* follow a specific morphophonological rule: They are pronounced as voiced phonemes /z/ and /d/ after stems ending in a vowel, a voiced consonant, or a diphthong or triphthong (e.g., *spilled* or *bags*) and as voiceless phonemes /s/ and /t/ after a stem with a final voiceless consonant (e.g., *hearts* or *missed*). After sonorants, *-(e)s* is pronounced as /ɪz/ (e.g., *busses*), and *-(e)d* is pronounced as /ɪd/ after stem-final alveolar phonemes (e.g., *waited*) (Dušková et al., 2006). Using these rules for regular verbal and nominal inflection gives us an opportunity to create novel English verbs to which we can add bound inflectional morphemes with and without potential morphosyntactic information (e.g., /vɪlz/ and /naɪld/ in which the voiced allomorph follows a voiced stem-final phoneme and is, therefore, morphological or /vɪls/ and /naɪlt/ in which a voiceless allomorph follows a voiced stem-final phoneme and is, therefore, non-morphological) and use them for the investigation of regular inflection at the sublexical level. It is evident that especially English regular verbal inflection constitutes a suitable framework for looking closely at morphology storage since verbal inflection represents a linguistic space in which morphological and phonological processing co-occur (as the phonological processes described above can be resisted in the presence of inflectional morphology, and the use of morphosyntactic rules enables to inscribe morphological information into novel words).

English irregular inflectional morphology appears in a number of word classes. In nouns, there are instances of phonetic spirant alternations (*bath [baːθ] – baths [baːðz]*), vowel mutation (*woman – women*), the use of the *-en* suffix (*ox – oxen*), zero plurals (*sheep – sheep*), or suppletive forms (i.e., using a different lexical form altogether; *Madam – ladies*). Instances of vowel mutation and suppletive forms also appear with adjectives and adverbs (e.g., *far – further*, *little – less, well – better*) and verbs (e.g., *grow – grew, be – was*). On top of that, English irregular verbs also make use of the irregular *-t* suffix (*learn – learnt*), a consonantal stem change *(make – made*), and the suffix *-n* in past participles (*choose – chosen*) (Dušková et al., 2006). These forms are usually referred to as unproductive (e.g., Dušková et al., 2006); this claim has, however, been challenged by recent psycholinguistic studies that show the influence of analogy (both in terms of regular and irregular word forms) on the production of novel morphologically-inflected forms (see, for instance, Albright & Hayes, 2003; Ambridge, 2010; Blything et al., 2018). It should also be stressed that while some of these phenomena happen in word-final positions, some of them appear in other positions.

Czech bound inflectional morphemes occupy the word-final position as well; yet, in Czech, the morphological repertoire is much more varied. Similarly to English, it encompasses both declination and conjugation, i.e., the change of nominal and verbal word forms by means of inflection, respectively. In terms of declination, Czech uses inflectional morphemes to express (i) nominal features of number, gender, and case in nouns, adjectives (and participles), relative pronouns, and inflected numerals (e.g., *pán-ovi, mladé-ho, které-mu*[2]), (ii) number and case in past participles (e.g., *psal-a*[3]), and (iii) case in personal pronouns (e.g., *vám-i*[4]). In terms of conjugation, inflectional morphemes express verbal features of person and number (e.g., *vaří-m*[5]). Czech inflectional morphemes are generally bound to the right word-final position (e.g., *spal-a*[6]); only a postfix can follow (e.g., *to-mu-hle*[7]) (with the exception of specific pronominal forms in which the postfix is further followed; e.g., *které-mu-ž-to*[8]) (Osolsobě, 2017). Since Czech is a morphologically richer language, it has a greater repertoire of bound morphemes that express inflection, and it is, therefore, very interesting to compare Czech, which is morphologically rather rich, and English, which is in general much poorer in terms of inflectional morphemes with which it operates. Their comparison might undercover the influence of morphologically rich L1 on the acquisition of morphologically poorer L2. However, it should be stressed that the two languages will not be compared directly in this thesis; the thesis will focus rather on how speakers of a highly inflectional language deal with a less inflectional L2, and thus how representations from one language may influence processing in the other.

2.2 The acquisition of inflectional morphology

Previous experimental and longitudinal studies show that in order to fully acquire the inflectional morphology of their mother tongue, children have to pass several developmental stages. Findings from nativist studies indicate that these developmental stages are more or less universal for each L1 speaker. This led to the conclusion that the acquisition of inflectional morphology in children is predictable and follows an expectable sequence. In pioneering work, Brown (1973) surveyed the production development of children between 20 and 36 months of age and found the following sequence of how they acquire functional items:

1. progressive *-ing*
2. plural morpheme *-(e)s*

[2] *To a man, young, to whom*
[3] *She has written*
[4] *By you*
[5] *I'm cooking*
[6] *She slept*
[7] *To this one*
[8] *To whom*

3. *‘s*
4. determiners *the, a*
5. *-ed*
6. verbal *-s*
7. auxiliary *be*

The child’s early awareness that word endings have a prominent role (and they are subsequently essential for further language development) is evident from the list. Cruttenden (1979) proposes three stages of how inflectional morphology is acquired: (i) memorization of individual words, (ii) acquisition of regular default rules and their application to base forms (e.g., the addition of *-ed* for past tense and the subsequent creation of such forms as *runned* instead of *ran* – a process known as overgeneralization, see, for instance, Marcus et al., 1992), and (iii) the correct use of morphologically-inflected words.

Previous longitudinal studies have also shown that Italian inflectional morphology is usually mastered before free morphemes (see, for instance, D’Amico et al., 2001; or Leonard et al., 2002). Caprin and Guasti (2009) demonstrated the same finding, testing 59 Italian children in the 3rd age of life, using cross-sectional data. Similarly to Pizzutto and Caselli (1992), their data showed that the acquisition of verbal morphology surpassed free morphology in the quickness of acquisition (while verbal inflectional morphology was mastered, the child participants, for instance, omitted articles).

With her famous wug test, Berko (1958) provided experimental evidence that children are able to infer regular inflectional patterns from their native language and extend these patterns to new words. Berko used novel words to test how children acquire morphological rules, showing them pictures of imaginary creatures or activities and asking them to supply plurals of novel nouns (e.g., *wug*) or past forms of novel verbs (e.g., *rick*). She argued that if a child is capable of extending the regular morphological rule for plurals to novel words (e.g., from *witches* to *gutches*), they have already internalized the rules for creating plurals, without anyone else teaching them (Rosenbaum, 2011). Berko was the first to show that pre-schoolers infer generalizable rules from the language and do not rely on memorizing stems as units together with the inflectional morphemes (Karmiloff & Karmiloff-Smith, 2001).

With children having the ability to infer inflectional rules directly from the language, we can explain how they are able to generate regular inflected forms of novel words. This, however, does not explain their production of irregular forms. It is sometimes believed that while regular forms are generated through the use of a default rule, irregular forms are learned by rote and retrieved directly from our memory (Pinker & Ullman, 2002). However, this does not explain why some of the irregular forms form similarity groups (e.g., *put – put, cut – cut*), nor why children are able to extend these analogical patterns to novel words. Ambridge (2010), for instance, showed that children aged 6 to 10 favourably judged novel irregular past forms that phonologically resembled existing irregular verbs. Similarly, Albright and Hayes (2003) and Blything et al. (2018) identically reported that their participants (adult English native speakers and English children from 3 up to 10 years of age, respectively) were "more likely to produce and favourably judge" (Blything et al., 2018: 3) irregular forms of these novel words that resembled existing irregular verbs. All these findings provide evidence for the productivity of irregular patterns, contrary to some previous views that posited that irregular forms are not productively applicable to new linguistic items.

When describing the acquisition of irregular morphologically-inflected forms, previous studies have shown that monolingual children pass the following stages on their way to the proper acquisition of irregulars: First, correct irregular form is inferred from the language and used. Then, a regular rule is spotted and applied to all morphologically-inflected forms by an extension (even the irregular ones) since "children are pattern makers" (Clark, 1987: 19). Children then undergo a period in which both overregularized and correct irregular forms are used. Bowerman (1982: 342) notices that "irregular forms rarely drop out, but rather continue to compete with their overregularised counterparts throughout the period of error making." This competition might last for "a period of months to years" (Maratsos, 1987: 19) until the child starts to use the irregular form correctly again and it becomes more firmly rooted in their language system (Avram, 2002). In the end, both regular and irregular forms are used correctly. Interestingly enough, certain inflections hardly ever get overgeneralized by children – one of the examples would be verbal *-ing* (Avram, 2002), the first acquired inflection by monolinguals according to Brown (1973).

Empirical data of spontaneous child language also show that, apart from overgeneralization of the regular pattern, children also undergo a period in which they incorrectly apply the irregular pattern to regular forms in a process known as irregularization (e.g., Xu & Pinker, 1995). Xu and Pinker (1995: 534) notice that "[other] children, beginning with the regular rule, abandon it for an irregular rule which they indiscriminately apply to all verbs, only later separating out the truly irregular ones and returning to the regular rule for the rest." Avram (2002) offers an explanation of this being the result of a retrieval failure, in which the phonological pattern of similar forms overlap and children produce the irregular form based on analogy to other irregular forms stored in their lexicon. It is, however, essential to mention that in comparison to overgeneralization, the rate of irregularization is rather low, and most of these "errors seem to be sporadic malfunctions in a system designed to suppress them, not recurring products of the system" (Xu & Pinker, 1995: 553).

The repertoire of inflectional morphemes is comparatively richer in Czech, as English is a predominantly analytical language, in contrast to Czech being a synthetic one (Dušková et al., 2006). Given the number of morphological functions expressed by inflectional morphemes (expressing person, gender, number, case, tense, etc.), Czech children need to focus closely on the endings of words, in general, to infer the correct meaning of the word and consequently of the whole sentence. This has consequences on Czech children's production. Previous research done on Czech children shows that while toddlers might leave out some lexical syllables in their speech, they never omit word endings (Lopatová, 2012). Smolík (2002) has also shown that while the use of verb endings might be limited in Czech two-year-old, their morphological system prevents them from incorrect usage of inflected forms that they have already acquired and used correctly. Similarly to English, Czech children start their inflectional development by acquiring the 3rd person singular of the present-simple tense (Smolík, 2002). Similar findings are reported in toddlers learning inflectional languages.

Guasti (1993) showed that even very young Italian children (around the age of 1;11 to 2;1) possess the knowledge of a subject-verb agreement and use it correctly. The author exemplified this on the distinction between finite and non-finite verbs, mainly on the distribution of non-finite verbs (that are to be found after the main verb and after a preposition in Italian) and on the placement of clitics (that appear on the left position of finite verbs, but on the right position of non-finite verbs). The data showed that the participants proved their awareness of verbal agreement morphology, purposefully distinguishing between finite and non-finite verbs. Moreover, the participants used verbal forms with several different bound morphemes, which points to the fact that they were already aware of verbal inflections, and verbal forms, therefore, could not be learned as units but were decomposed into stems and bound morphemes. Guasti's findings of the presence of at least some functional categories in the child's early grammar are also supported by previous research (see, for instance, Hyams, 1992; or Poeppel & Wexler, 1993). Similarly to Smolík (2002), Guasti (1993) showed that Italian children rarely show incorrect use of verbal agreement morphology in their production.

In contrast, speakers of morphologically poorer languages tend to acquire proper inflectional morphology later than speakers of morphologically richer languages (Guasti, 2017).

In their studies on the acquisition of mini-paradigms[9], Bittner et al. (2011) showed that children of different L1s (e.g., Spanish, German, Finnish, Italian) master several inflectional morphemes early on; yet, English-speaking children behave differently and only an early mastery of the *-ing* morpheme has been reported while the *-s* and *-d* morphemes require more time to be acquired fully. Guasti (1993; 2017) ascribes this discrepancy to the difference in morphological richness between English and the other tested languages. The reasons for such difficulties in children with a poorly inflected L1 are twofold: morphologically rich languages tend to realize their inflectional morphemes phonologically and prosodically more prominently (compare, for instance, the difference between adding a new syllable in Czech *jí-me* and the lack of any inflectional morpheme in English *we eat-0*) and English-speaking children simultaneously are not exposed to inflectional morphemes to the same degrees as speakers of a morphologically rich L1 (which usually inflect words for person, gender, case, etc.). Paradigms, therefore, seem to be acquired later in speakers of morphologically poorer languages than in speakers whose L1 is a highly-inflectional language (Cilibrasi & Tsimpli, 2020). The fact that Italian children, whose L1 operates with a variety of inflectional morphemes, start mastering inflectional paradigms sooner than their English peers, whose L1 morphology is considerably poorer, has been manifested in previous research as well (Hyams, 1992; Pizzuto & Caselli, 1992; or Guasti, 1993).

Historically, Brown's study (1973) provided the first evidence that children acquire their L1 grammar in a natural order. A question, therefore, arose as to whether the same could also be said about second-language acquisition. Dulay and Burt (1973), for instance, attempted to explore the morphological order of acquisition in 151 Spanish children with English as L2 and found that the acquisition order of inflections was similar for all of their participants. Dulay and Burt (1974) then explored if the learner's L1 is capable of influencing the order of morphological acquisition in any way, recruiting children with L1 Chinese and Spanish. Both tested groups showed virtually the same order of acquisition, regardless of their L1 or its (dis)similarity with L2 English. These findings provided some evidence for the belief that the universal order of morphological acquisition can also be extended to children second-language

[9] The so-called mini-paradigms describe inflectional productivity, characterized by the emergence and repeated occurrence of phonologically distinct morphologically-inflected forms of the same lemma in spontaneous production (Bittner et al., 2011).

learners with no or minimal effect of L1. Some evidence against this was proposed by Hakuta (1976), who presented a longitudinal observation of a Japanese girl learning English as her L2, who showed deviation from the order of English native speakers. Hakuta proposed that the observed performance might be explained by grammatical differences between English and Japanese (more specifically, the absence of a feature in L1 Japanese resulted in its later acquisition in English), and he suggested that first language needed to be taken into account when investigating the order of morpheme acquisition.

A subsequent question arises as to whether the same also applies to adult L2 learners. Bailey, Madden, and Krashen (1974) recruited adult L2 learners of different L1s, exposure to English, and formal instruction and tested their knowledge of morphology. The authors found that the acquisition order of L2 inflections was comparable among the participants and similar to previous findings found in children L2 learners, irrespective of the form of instruction, previous exposure to English, or L1. After that, the universal order of morpheme acquisition has been widely accepted in literature. However, building on the work of Hakuta (1976) and others, more recent studies on the role of L1 in morpheme acquisition seem to suggest that L1 might be influencing linguistic behaviour. Luk and Shirai (2009) investigated previous research on English speakers of Chinese, Korean, Japanese, and Spanish and their acquisition of English morphemes. The authors found that the language background seems to influence the acquisition of inflectional morphemes. The order of acquisition might, therefore, not be fully explainable of L2 linguistic behaviour, and individual learner variability in L2 morphological production and perception should be studied instead.

Compared to these pioneering studies of the 1970s, research done in the 1980s and 1990s seems to move on to identifying the variables that influence language acquisition and cause performance variance. Learner's L1 was identified as one of these fundamental factors (see, for instance, Hakuta, 1976 above), and this finding led to further investigation in a multiple-determinant approach (Goldschneider & DeKeyser, 2001; Kwon, 2005; Gass & Selinker, 2013). This line of research attempted to explain the universality of morpheme acquisition through the interaction of several possible factors, which seems to better predict learner's L2 variance, giving it a new perspective. Factors investigated under this approach included perceptual salience, morphophonological regularity, syntactic complexity, frequency, semantic complexity, native language transfer, individual variance, and levels of morpheme activation (Kwon, 2005). Goldschneider and DeKeyser (2001) imported these factors into a multiple regression analysis, and their findings show that "a very large portion of the total variance in the acquisition order of grammatical morphemes in English can be explained by a combination of five factors: perceptual salience, semantic complexity, morphophonological regularity, syntactic category, and input frequency" (Kwon, 2005: 17). Therefore, it seems that, given the complexity of (second-)language learning, further studies will need to consider these and other factors to explain the order of L2 development.

2.3 L2 challenges in morphology acquisition

Although some previous studies attempted to find evidence for the similarity of L1 and L2 acquisition (e.g., Cunnings, 2017), the majority of contemporary second-language studies focus on the problems that an L2 learner generally faces when acquiring their second language even after years of active exposure. While case studies might not be a theoretical foundation for this research, some individual stories give us a sense of the topic and can be useful in explaining the type of problems that motivate research like the current one. For example, Long (1997) described the linguistic behaviour of his subject Ayako, whose accuracy of past-tense production hardly amounted to 50 % even after 52 years of having been exposed to English. Lardiere (1998a, 1998b) shows a similar finding, with her Chinese subject Patty that used the inflectional morpheme *-ed* correctly only in 34 % cases even after having been exposed to English for 18 years. L2 difficulties in acquiring English inflectional morphology are well-documented in second-language literature (see, for instance, White, 2003; Jiang, 2004; or Clahsen et al., 2010). The acquisition and processing of morphosyntactic information expressed by inflectional morphemes have been proved by research as one of the most challenging tasks for the L2 learner (see, for instance, DeKeyser, 2005), fuelling further research of the specifics of L2 acquisition of (inflectional) morphology.

There has been extensive effort to explain the L2 inconsistencies in using L2 morphology and the challenges L2 learners confront. One line of research (e.g., Eubank et al., 1997; Hawkins & Liszka, 2003) generally proposes that the non-target-like linguistic behaviour is caused by "an underlying syntactic impairment in the domain of functional categories or features" (Solt et al., 2003: 1-2), while another line of research proposes that it is not the second-language syntactic representations that would be impaired, but the learner's struggles lie outside the scope of syntax (e.g., Lardiere, 1998; Prévost & White, 2000). For instance, for her subject Patty, Lardiere (1998) proposed an issue of "mapping" between the abstractness of the past tense and its realization (surface) form. Lardiere (1998) showed that phonological factors (for instance, L1 interference effects) might hinder L2 morphosyntactic production and perception. Her Chinese subject Patty frequently omitted *-t/-d* morphemes in the clusters at the end of the word, not only in inflected past-tense forms but also monomorphemically. Lardiere used this finding as evidence that the failure to supply correct past-tense forms cannot be explained by the absence of such a feature [±past] in Patty's L1, but it might be a consequence of a Chinese constraint against final consonant clusters (Bliss, 2006). Such a role of phonology has been investigated in connection to the consonant cluster reduction hypothesis. Bliss (2006) presented similar findings for Chinese speakers and their failure to provide correct forms of English plural morphology. Similarly to Lardiere, Prévost and White (2000) showed that their French and German participants did not exhibit impaired syntactic representations but rather struggled with mapping the abstract grammatical feature onto its realization surface (see the missing surface inflection hypothesis below).

Several other theoretical hypotheses have, therefore, proposed challenges with inflectional morphology that second-language learners experience. An important theory is the so-called "prosodic transfer hypothesis." In an influential publication, Goad, White, and Steele (2003) suggested that even L1 prosodic representations might be at the base of the incorrect L2 production of inflectional morphology, arguing that Chinese L2 speakers might struggle with past-tense inflections due to the difference between English and Mandarin prosodic structures. Yet another theory is the failed functional features hypothesis which attributes the learner's struggle with L2 inflectional morphology to morphosyntactic constraints in their L1 (see, for instance, Hawkins & Chan, 1997). With the lack of these morphosyntactic categories in L1, the same categories will not get activated, and are, therefore, inaccessible in their L2 production and thus virtually unacquirable. Chinese incorrect production of English past tense could be then explained by the lack of the past feature in the learner's L2 linguistic representations (see, for instance, Hawkins & Liszka, 2003).

Clahsen and Felser (2006) then proposed the shallow-structure hypothesis that suggests that L2 processing relies more on non-grammatical information sources than the processing of native speakers due to the exceeded period for grammar acquisition (Bosch et al., 2019). The theory predicts that L2 learners struggle with rule-based knowledge and rely more heavily on declarative memory and rote storage (Friedline, 2011). The missing surface inflection hypothesis (Haznedar & Schwartz, 1997) then holds that L2 learner's grammar is not impaired and "that abstract properties may be present in the interlanguage grammar without being systematically realized morphologically" (Oshodi, 2014: 19). Problems with accessing and mapping morphology are also present. The underlying representation might, thus, be correct, but the surface structure is not native-like.[10] All these theoretical approaches attempt to bring an explanation to the apparent challenge that L2 morphology represents for L2 learners, yet no clear answer has been given so far as to which of the hypotheses provides the correct interpretation of the L2 variability. One thing remains clear, though: L2 learners seem to struggle with the acquisition of inflectional morphology, and specific steps need to be taken to lessen their burden in L2 classrooms.

2.4 Models of language processing

In connection to how L2 learners acquire inflectional morphology, two approaches are usually proposed as direct opponents: (i) the single-route model (Halle & Mohanan, 1985; Bybee, 1995), which posits that all morphologically inflected forms are processed through one mechanism, and (ii) the dual-route model (Prasada & Pinker, 1993; Chialant & Caramazza, 1995; Clahsen, 1999), which posits that regular and irregular forms are processed by two different mechanisms. Sometimes, a third "hybrid" mechanism of morphological processing is quoted (e.g., Albright & Hayes, 2003), which "makes use of multiple, stochastic rules" (Albright & Hayes, 2003: 120). Special attention is being paid to the English past tense since English encompasses a relatively clear-cut distinction between the two forms and, therefore, suits perfectly to the purposes of deciding between these two approaches.

In terms of production, the dual-route model, sometimes also called a words-and-rules model (Westermann & Ruh, 2012), posits that regular and irregular morphologically-inflected forms are produced by two different mechanisms: The regular forms are produced by a default rule (in case of past tense, this means adding -*ed* to the stem of a verb, e.g., *jump/jumped*), while irregular forms are stored (and during production retrieved) as memorized units from our associative lexical memory (e.g., Pinker, 1991; Pinker & Ullman, 2002). As Pinker and Ullman (2002: 456) explain, "[a] stored inflected form of a verb blocks the application of the rule to that verb (e.g., *brought* pre-empts *bringed*); [e]lsewhere (by default) the rule applies […]" In practice, this means that speakers apply the rule "whenever an irregular form fails to be either

[10] In line with the shallow structure and missing surface inflection hypotheses, this thesis also assumes that any potential non-target performance of Czech L2 learners may not be indicative of the abstract properties being absent in their interlanguage. Our participants, especially the lower-proficiency ones, might be also more prone to the use of declarative memory and rote storage, both in production and perception of inflectional morphology.

retrieved directly from memory or generated by analogy with stored irregulars" (Blything et al., 2018: 3).

However neat such a model may seem, its claims have been somewhat undermined by recent empirical findings suggesting that regular forms are also capable of showing frequency and family-resemblance effects characteristic of associative storage (e.g., Alegre & Gordon, 1999). Such evidence seems to support more the single-model approach (see, for instance, Rumelhart & McClelland, 1986), in which – when talking about past-tense morphology – "all past-tense forms are generated by the same associative mechanism" (Westermann & Ruh, 2012: 650). As Pinker and Ullman (2002: 457) explain, "[t]he key to these pattern associators is that rather than linking a word to a word stored in memory, they link sounds to sounds. Because similar words share sounds, their representations are partly superimposed, and any association formed to one is automatically generalized to the others. This allows such models to acquire families of similar forms more easily than arbitrary sets, and to generalize the patterns to new similar words." According to the supporters of the single-route model, this principle is key in past-tense production. If a past-tense form is not evoked straight from the mental lexicon, "the past-tense form is generated on the basis of analogy to phonologically similar verbs regardless of their regularity" (Blything et al., 2018: 2). A child trained on the change of *fling* and *cling* into *flung* and *clung* may generalize the past form of *spling* as *splung*, just as they may generalize *plip* as *plipped* based on the existing pairs of *clip-clipped* and *flip-flipped* (Blything et al., 2018). Generalization thus underlies the production of both past-tense forms. Bybee and Slobin (1982: 266) define such generalization as schemas that describe "general phonological properties of a morphological class and [are] used in organizing and accessing the lexicon."

As for novel verbs, "the single-route model predicts that the likelihood of a novel verb being produced in regular past-tense form is positively associated with [their] phonological similarity to existing regular verbs" (Blything et al., 2018: 3), just as for the irregular production, such a form is "positively associated with its phonological similarity to existing irregular verbs" (Blything et al., 2018: 2) based on analogy. It is, therefore, evident that both models assume similar scenarios for the generation of irregular past forms, which are "stored in associative memory and used in analogical generalization" (Blything et al., 2018: 2). Investigating irregular verbs is, therefore, not helpful in deciding which model is better at describing morphological productivity of past-tense inflection, and the production of regular inflection needs to be studied instead. For regular production, the single-route model posits that the more similar to an existing regular verb a given novel word is, the higher the probability of its being produced in a regular way, while the dual-route model posits that our mind first searches our mental lexicon for the stored irregular form, and if such a retrieval fails, a regular rule of adding *-ed* to the stem will be applied by default.

In terms of perception, the single-route mechanism states that all morphologically-inflected forms "are processed with the same system" that "generates inflected verbs based on regularities in the input" (Cilibrasi et al., 2019: 749). The dual-route model then posits that regular and irregular inflected forms are processed by two different systems. Pinker & Ullman (2002: 464) suggest that while "irregular forms are stored in the lexicon, a division of declarative memory, regular forms can be computed by a concatenation rule, which requires the procedural system." This means that regular forms are processed through decomposition in which the stem and the bound inflectional morpheme are separated, and the word is parsed (Cilibrasi et al., 2019). Since the perception chapter focuses mainly on regularly inflected forms, it is essential to highlight the difference applicable to these forms as well – they are processed either (a) within one analogical system where the inflected forms are stored as whole-word forms linked by their formal and semantic similarities to other words (see, for instance, Bybee, 1995; or Rumelhart & McClelland, 1986) or (b) within a combination of two systems, where the morphologically-inflected forms can be processed both associatively (through the whole-word storage) and by rules (through decomposition and parsing) (see for instance Chialant & Caramazza, 1995; Clahsen, 1999; Pinker, 1991).

Previous research on L2 inflectional morphology has not been able to give a clear answer on which of these models best describes second-language performance. Quite recently, attention has also been paid to the possible redundancy of our language system, suggesting that both mechanisms could be activated in parallel and co-exist. Schreuder et al. (1999), for instance, propose a redundant model of morphological processing, in which two types of analyses operate simultaneously and in a complementary way: single- (route-based parsing) and dual-route mechanisms (rule-based parsing). These analyses operate in parallel, but each system has varying importance connected to how frequent the presented item is in the language: frequent items (or, when speaking of novel words, items with high phonotactic probability) are "to be parsed with the rote system, while infrequent items [or, when speaking of novel words, items with low phonotactic probability] tend to be parsed with the rule system" (Cilibrasi et al., 2019: 762). This view is further supported by observed frequency effects in the production of regularly inflected verbs (see, for instance, Alegre & Gordon, 1999; or more information on the redundancy models in production, see Chapter 4 below). Albright and Hayes (2003: 120) then argue for a new (third) "hybrid" mechanism of morphological processing, which "makes use of multiple, stochastic rules," combining associative processes and phonological micro-rules into structural analogy. Despite the descriptive power of such a redundant proposal, these claims are inconsistent with some influential views on language processing. Particularly, the hypothesis that both models might be at play in language perception would be a clear contradiction to the models that rely heavily on the concept of economy, such as Chomsky's minimalist program (Chomsky, 1995), since a redundant proposal is not, in principle, minimal.

2.5 An overview of L2 production, perception, and interference effects[11]

In English, word endings receive particular attention since they are involved in a number of morphological processes. Special attention is paid to verbal inflection, which requires both phonological and morphological processing and, therefore, is a suitable ground for investigating morphology. The perception of morphologically-inflected forms has been a ground for a debate between those who suggest that regular morphologically-inflected forms are stored as wholes (in which the stem is stored together with the bound morpheme) in our lexicon (also known as the atomist theory represented, for instance, by Stemberger & MacWhinney, 1986; Bertram et al., 2000; Tomasello, 2006) and those who suggest that regular morphologically-inflected forms are stored separately as stems and bound morphemes (also known as the decompositional theory in which morpheme stripping plays a central role; see, for instance, Berko, 1958; Pinker & Ullman, 2002a; Post et al., 2008; Guasti, 2017). Previous research shows that morphological decomposition might be active also in the sublexicon (e.g., Grainger and Ziegler, 2011; or Cilibrasi, 2016), suggesting that in perception, English native speakers are sensitive to the presence or absence of morphosyntactic information in novel words with no access to the word's semantics (Cilibrasi, 2016). Previous studies on (more proficient) L2 learners (e.g., Jiránková, 2017) have shown that even non-native speakers of English might be sensitive to this type of morphological information and decompose morphologically inflected words into stems and affixes. This dissertation, therefore, aims to describe how/if inflections are stored in the mental lexicon of L2 learners of all levels and whether this storage shows any development during the L2 acquisition.

[11] This section presents only a very brief overview of the perception, production, and transfer theories and previous studies. For more details, see individual chapters focused on perception (Chapter 3), production (Chapter 4), or transfer effects (Chapter 5).

Mirroring the debate on perception, the debate about how native speakers produce novel morphologically-inflected forms attributes the productivity either to the application of rules (e.g., Prasada & Pinker, 1993) or to analogy based on stored exemplary forms (e.g., Bybee & Slobin, 1982). For its relatively clear distinction between regular and irregular patterns of inflection, English past-tense morphology provides a suitable ground for deciding between these two approaches. The two models of morphological processing discussed above can be, therefore, adapted for past-tense production: the single-route model (e.g., Bybee & Moder, 1983) claims that both regular and irregular past-tense forms are produced by analogy based on forms already stored in our lexicon and the dual-route model (e.g., Prasada & Pinker, 1993) claims that regular past-tense forms are produced through the application of a rule *–ed*, while irregulars are stored as units and used in analogical generalization (Blything et al., 2018). This dissertation thus aims to describe which production mechanism is used by L2 learners and whether this mechanism changes during L2 acquisition.

When it comes to transfer effects, recent findings show that L1 does not seem to be easily separated from the processing in L2 (e.g., Kroll et al., 2002). Research on transfer effects generally makes a distinction between the word association model with a direct mediation of L1 and the concept mediation model with no mediation of L1 (Potter et al., 1984). Previous studies show that L2 learners seem to decrease their dependence on L1 with higher proficiency and access L2 words directly (Kroll et al., 2002). In connection to the relation between the two lexica and their storage, van Heuven et al. (1998) proposed a distinction between language-dependent lexica with one integrated lexicon for the two languages and language-independent lexica, in which each language has its own lexicon. Concerning the online processing of words from the two lexica, a difference is often made between a language non-selective access, in which "words of both languages are activated (or considered as candidates)" (Van Heuven et al., 1998: 458), and a language selective access, in which "only words belonging to the targeted language" get activated (Van Heuven et al., 1998: 459). Several research findings show that during language processing, L2 learners seem to activate the lexica of both languages in non-selective language access (Van Heuven et al., 1998; Lemhöfer & Dijkstra, 2004; Brenders et al., 2011). Previous studies have also shown that proficiency is a decisive factor in the organization of the two lexica, with higher interference in less proficient L2 learners (see, for instance, Kroll, 1993; Talamas et al., 1999; or Harrington, 2006). This led us to investigate the effect of L1 on L2 learners of English at different proficiency stages. As a morphologically rich language with a substantial number of inflectional suffixes, Czech provides an interesting comparison with the morphologically poorer English inflectional system. The mismatch between the two systems can bring more insight into the role of L1 on the acquisition of L2 morphology and uncover its strengths and weaknesses. While Czech might perform as a facilitator of the acquisition and promote the learner's performance due to the higher morphological awareness productive in their L1 (see, for instance, Ku & Andersson, 2003), it might also hinder the acquisition of L2 English morphology and clash with it.

2.6 Hypotheses

Given the complex bundle of research presented in the previous pages, we hypothesize that L2 learners might be sensitive to the presence of potential morphosyntactic information and perform morpheme stripping during perception, while some lexical effects might be present as well (e.g., Jiránková, 2017). In terms of production, we operate with a tentative hypothesis that lower-proficiency learners might be more inclined to produce past-tense forms of verbs rules using rules, while higher-proficiency learners might be more prone to the use of analogy based on exemplars they have stored so far. Regarding transfer effects, Czech might function as a facilitator of the acquisition and promote the learner's performance due to the morphological richness of their L1 and their subsequently heightened sensitivity to inflectional morphology in general (see, for instance, Ku & Andersson, 2003), or it might also hinder the acquisition, leading to interference in representations.

2.7 Steps taken and their outcomes

In order to test the use of a generalization mechanism and not activate semantics, it is generally agreed to employ novel verbs in the testing (e.g., Prasada & Pinker, 1993, Albright & Hayes, 2003; Ambridge, 2010; Cilibrasi, 2016; Jiránková, 2017; Blything et al., 2018). The use of novel verbs will enable us to focus solely on the word form without access to the lexical meaning of the word, investigating the implicit sublexical knowledge of L2 learners. All experiments in this dissertation, therefore, use novel words that were deemed phonotactically legal in English to test the perception and production of inflectional morphology in second-language learners of English, most specifically to look at how inflectional morphemes are stored and whether morphologically-inflected forms are produced through rules or analogy. The aim of this dissertation is, therefore, to look more closely into the development of second-language inflectional morphology both in perception and production, find evidence for either the single- or dual-route model (or both), and describe the role L1 has in this development. To test our hypotheses, three experiments were carried out. The experiments and their findings are shortly summarised below.

1. The first experiment attempts to answer the question of whether L2 learners use morphological decomposition (also known as morpheme stripping) during language processing. In order to find evidence, we explored the perception of morphological contrasts applied to the endings of novel words. With this experiment, we wanted to understand whether L2 learners detect and process potential morphosyntactic information differently than native speakers when presented with sublexical stimuli with no access to the lexicon. In order to do so, we focused on whether novel words with and without any morphosyntactic information are discriminated differently. This experiment extends my previous research on word-ending perception in second-language learners of English, in which 60 L2 learners of English at the B1, B2, and C1 levels of proficiency were tested. For this dissertation, another 31 A1- and A2-level L2 learners and 4 A0 learners were recruited. We conducted a reaction-times experiment in which the subjects were supposed to decide whether the items in the presented minimal pairs of novel words were identical or different. The results suggest that L2 learners might be sensitive to the presence of morphosyntactic information even in the absence of semantics, and they might use morphological decomposition. The results also suggest that parallel to morpheme stripping, some frequency effects might also be active and similarly effective in predicting the performance.

   To see if the findings also apply to other languages, the study was compared to Cilibrasi's study (2016), which tested 30 adult native speakers of English. The author's findings suggest that monolingual speakers are also sensitive to the presence or absence of morphosyntactic information at the sublexical level, and morphological decomposition is likely used. These findings suggest L2 learners process English inflectional morphology similarly to native speakers. The inclusion of the data collected from L2 learners at the lowest proficiency level allows us to focus our attention also on cross-language interference effects. Given that our participants preserved a performance pattern comparable to that of other L2 learners and native speakers, the morphological richness of Czech might be the explanation of these findings, and Czech might be influencing the performance.

2. The second experiment attempts to answer the question of whether L2 learners use rules or analogy when producing past-tense forms and whether this strategy changes during the course of language acquisition. In order to inspect this issue, we tested how L2 learners generate novel morphologically-inflected forms, using an elicitation task to examine which of the two previously-mentioned models best accounts for L2 learners' morphological productivity. 88 adult English L2 learners at A1-C1 proficiency levels and a control group of 9 native speakers heard sentences about someone performing a novel activity (e.g., *The baby likes to dize. Look, there he is dizing. Everyday he dizes. So yesterday he…*). Past-tense forms were then elicited by asking the participant to say what the agent "did yesterday." Produced forms and reaction times were recorded. The results on the produced forms indicate that the L2 acquisition of the English past-tense is characterized by a progressive development from the mastery of mechanistic rules at the A1, A2, and B1 levels to the refinement of their application when the L2 learners eventually start spotting analogical patterns of existing verbs at the B2 and C1 levels. The analysis of the reaction times also suggests that in the case of A1 and A2 levels, Czech has a facilitatory effect and positively influences the performance.

   To see if the findings once again apply to other languages as well, L2 learners were compared to native speakers. The findings show that the second-language speakers show developmental changes in the inflectional morphology that come closer to adult native speakers with the higher proficiency of B2 and C1 levels since the native speakers used analogy exclusively.

3. In the third experiment, we aimed to look at any potential cross-language interference effects to see if Czech could be used as a facilitator in L2 performance. In order to do that, we recruited 30 participants that created a continuum of language proficiency and presented them with a lexical decision task. The participants were asked to decide whether a series of items they heard were real English words or not. The stimuli consisted of real English words and novel words, which were divided into two halves: one phonotactically legal both in English and Czech and one phonotactically legal in English but not in Czech. This was done in an attempt to understand any possible transfer effects from L1 Czech. With this design, the experiment also attempted to

answer whether the lexicon access is (non)selective or language (in)dependent. The results of reaction-times and accuracy analyses showed that a mild facilitatory effect of L1 could be found not only at the lowest language levels but also in more proficient L2 speakers.

A schematic summary of the experiments is provided below:

Experiment 1: Perception experiment
(morphological vs. non-morphological novel words – stored as units or decomposition)

Tested on L2 learners (A0 – C1) and English native speakers (Cilibrasi, 2016)

Experiment 2: Production experiment
(single- vs. dual-route model – rules or analogy)

Tested on L2 learners (A1 – C1) and English native speakers

Experiment 3: Lexical decision experiment
(cross-language interference effects – the effect of Czech)

Tested on L2 learners (proficiency continuum)

With these experiments, we have shown that, at the sublexical level, the storage of inflectional morphology in L2 learners might be characterized by morphological decomposition, with some additional frequency effects. We also showed that the L2 acquisition of the English past tense is characterized by a progressive development from the mastery of mechanistic rules to the consequent application of analogical patterns based on existing verbs stored in our mental lexicon. Finally, the analysis of transfer effects suggests that Czech has some facilitatory influence, with novel words legal in both Czech and English being recognized more quickly. Since these findings offer not only theoretical contributions but also pedagogical ones, the dissertation will also discuss the potential consequences of our findings for L2 teaching methodology.

3. EXPERIMENT 1: THE PERCEPTION EXPERIMENT[12]

3.1 Introduction

The way how inflectional morphology is represented in the mental lexicon is the object of a long and intense debate in the psycholinguistic community. A significant number of studies have concentrated on how morphologically-inflected forms are stored in our minds and how the listener accesses them during language perception. Special attention has been paid to regular verbal inflection, which provides a particularly suitable framework for examining both the access route (i.e., how morphologically-inflected forms are processed) and the storage of inflectional morphology[13].

The processing of regular morphologically-inflected verbs in English, particularly, has been strongly debated since the advent of psycholinguistic research. One line of studies (e.g., Stemberger & MacWhinney, 1986; Bertram et al., 2000; Tomasello, 2000) proposes that regular morphologically-inflected items are stored as whole-word representations (i.e., units consisting of both the stem and the bound morpheme) in the mental lexicon (and accessed as such during perception). This hypothesis has been previously supported by the findings of lexical decision tasks in which participants showed strong effects of word frequency on reaction times (with quicker reaction times for more frequent items), which suggests that they must have retrieved the inflected form as a unit (Cilibrasi, 2016). The other line of research (e.g., Berko, 1958; Pinker & Ullman, 2002; Guasti, 2004; Post et al., 2008) proposes that regular morphologically-inflected items are stored as stems and affixes, not as whole units, but as separate items. During language perception, such inflected words would be divided into two parts in a process called morpheme stripping (Cilibrasi, 2016). As described in the theoretical chapter, regular English verbal inflection constitutes a suitable framework for distinguishing between the two proposed lines of research and for investigating word-position effects since verbal inflection represents a linguistic space in which morphological and phonological

---

[12] Parts of this chapter will be published (with some changes) in the *Journal of Monolingual and Bilingual Speech*. The paper was in press at the time of the thesis submission.
[13] Regular verbal morphology follows specific morphophonological rules and their perception might help us decide how morphologically-inflected forms are processed online and stored in our mental lexicon

processing co-occur. The word-final positions are special in that such phonological processes as deletion or assimilation occur only when no morphosyntactic information is present – if it is, even word-final positions can avoid the above-mentioned phonological processes. Also, following the English morphosyntactic rules in word-final positions enables us to inscribe morphological information into novel words.

Recent studies show that morpheme stripping may not be taking place only in the processing of lexical items but also sublexically in the processing of novel words without the influence of semantics. This idea has been previously proposed by Grainger and Ziegler (2011) for reading and by Cilibrasi (2016) and Cilibrasi et al. (2019) for first language perception. This study builds on previous research findings by Cilibrasi (2016) and Jiránková (2017), who investigated morpheme stripping in English native speakers (Cilibrasi, 2016) and more proficient second-language learners of English (Jiránková, 2017), testing items with no access to the lexicon (i.e., sublexical items, also known as novel words). These studies aimed to investigate whether stems and affixes of morphologically-inflected sublexical items are processed separately or together as one item (Jiránková, 2017). The perception experiment in this thesis then aims at describing the processes underlying the sublexical perception of verbal endings in L2 learners, including lower proficiencies. Since Czech is a language with rich inflection, it creates an interesting contrast with the notoriously poor inflectional system of English and, therefore, provides a suitable ground for investigating possible transfer effects of Czech on the acquisition of English inflectional morphology.

3.2 Theoretical background and hypotheses

3.2.1 Theoretical background

3.2.1.1 Atomist view

While a number of studies have already explored regular inflectional morphology, its generation, and storage, we are still faced with a number of unanswered questions. So far, the studies investigating how regular morphologically-inflected verbs in English are processed have reported findings that could be divided into two distinct lines – the atomist view and the decomposition view. The studies promoting the atomist view (e.g., Bertram et al., 2000;

Stemberger & MacWhinney, 1986; Tomasello, 2000) suggest that morphologically-inflected forms are stored as whole units in the mental lexicon and processed as such during perception. This means that, for instance, regular past-tense forms are stored together with their bound morpheme *-ed* (not separately as the stem and the past-tense inflection). Evidence for this hypothesis can be found firstly in studies using lexical decision tasks, which showed that participants are more accurate in producing regularly inflected words with high frequency (e.g., *smiled*) than regularly-inflected words with low frequency (e.g., *ruminated*) (Stemberger & McWhinney, 1986), suggesting that not only irregular forms but also highly-frequent regular words must be stored as units in our mental lexicon (however, this also means that lower-frequency regular forms, which include most English words, are not stored as units).

Secondly, relevant evidence can also be found in language-acquisition research (e.g., Tomasello, 2000) which suggests that children start storing words as units together with their inflections before moving on to morpheme stripping. In this view, children seem to store all morphologically-inflected forms as wholes (i.e., stems with the bound morpheme) in their mental lexicons. Only later, when they realize the principles for creating past inflected forms, they start extending the rule onto other words by analogy. The evidence for this claim lies in studies that show that the child's development of grammar initiates with lexically-specific expressions (Diessel, 2013). Diessel (2013: 5) exemplifies these early item-based constructions by the earliest child production, which consists mainly of "isolated words and holophrases" (e.g., *all-gone*), usually closely connected to a specific communicative function (Tomasello, 2000). The child eventually learns to decompose these constructions and analogically apply the rules to other words.

Further evidence for the atomist view can be found in studies with English phonotactic probabilities and their significant effect on reaction times. For instance, Jiránková (2017) showed that her participants used novel verbs' phonotactic probabilities to distinguish morphosyntactic pairs with identical and different items. Higher phonotactic probabilities amounted to quicker reaction times.

3.2.1.2 Decompositional view

The opposing line of research (e.g., Berko, 1958; Pinker & Ullman, 2002; Post et al., 2008) proposes that regular morphologically-inflected verbs are stored as stems and bound morphemes (or rules that govern them), separate from one another, and during perception, they are decomposed in a process called morpheme stripping (Cilibrasi, 2016). According to this view, the same process of combining the stem and the bound morpheme is used in the generalization of language forms during production. Pinker and Ullman (2002), for instance, suggest that regular past-tense forms consist of the verbal stem to which the rule for regular past-tense formation is added. This implies that only verbal stems are stored in our lexicon, along with specific sets of rules that are activated both during perception and production. This view of language processing is logically supported by the concept of language economy since it is more economical and less demanding for the memory load to store the stems and rules (bound morphemes) separately, rather than storing all forms individually (Cilibrasi, 2016). This advantage is not particularly evident in English, a language that operates only with a limited number of inflectional morphemes (e.g., *-s, -ed*, or *-ing*); however, the same cannot be said about highly inflected languages like Czech, where the economical aspect plays a much significant role given the number of inflections and the cognitive demand for which their storage would need to account.

Perceptual decomposition entails identifying inflectional bound morphemes and isolating them from stems in a process called "morpheme stripping." Previous psycholinguistic studies have shown that this process is by no means blind, i.e., that "the parser does not strip the morpheme without analysing the stem as well" (Cilibrasi, 2016: 105). Studies of the sublexical processing of language with no access to the word's lexical meaning, such as Caramazza et al. (1988), Clahsen (1999), and Post et al. (2008), have shown that "the analysis of the bound morpheme is synergistic to the analysis of the stem" (Cilibrasi, 2016: 105-106), i.e., that in order to analyse the affix, the participant first needs to analyse the final phoneme of the stem to examine whether the quality of the bound morpheme could be potentially morphological (Cilibrasi, 2016). Caramazza et al. (1988), Clahsen (1999), and Post et al. (2008) also show that morphological decomposition (or morpheme stripping) does not have to apply only lexically but also sublexically without access to the lexicon (even though, as Cilibrasi et al. (2019) note, the nature of this synergy is not fully understood yet). This notion of sublexical synergy is further supported by Grainger and Ziegler's (2011) study on sublexical processes in reading, by Cilibrasi (2016) and Cilibrasi et al. (2019)'s studies on the aural perception of inflected novel words in English native speakers, and by Jiránková (2017)'s thesis on the perception of more proficient L2 learners.

3.2.1.3 Previous studies

Up until the end of the century, studies explored mainly morphological processing of real words stored in the lexicon (e.g., Brown & McNeill, 1966; Browman, 1978; Nooteboom, 1981; Marslen-Wilson & Zwitserlood, 1989). Recently, studies of those effects in the sublexicon have been given more attention (e.g., Pitt & Samuel, 1995; Marshall & van der Lely, 2009; Grainger & Ziegler, 2011), investigating the perceptual level as well (e.g., Pitt & Samuel, 1995; Post et al., 2008; Cilibrasi, 2016; or Cilibrasi et al., 2019), though perception is, in general, less studied than production. Studies focusing on the sublexical perception in second-language learners are even rarer (e.g., Jiránková, 2017), and the presented research, therefore, has the potential to bring greater understanding of the processes underlying morphological perception.

Within the research on morphological processing, the most commonly studied area is verbal inflection since it represents a linguistic field where both morphology and phonology work closely together (phonological processes such as assimilation or deletion can be suppressed if morphosyntactic information is present in the word-final position). As mentioned above, the study of perceptual processing of morphologically-inflected forms in the sublexicon is still in its birth; however, exceptions to this exist: Grainger and Ziegler (2011), for instance, found out that their participants performed morpheme stripping, i.e., the decomposition into stems and bound morphemes, during reading, even in the sublexicon with no access to a lexical meaning. In their connectionist model of reading, the authors proposed that readers may detect bound inflectional morphology, such as *-s* or *-ed*, sooner than they access the word in the mental lexicon, suggesting that the participants work with "a fine-grained parser that detects grapheme chunks representing morphological information" (Cilibrasi, 2016: 189). This finding leads to the conclusion that the affixes are stored separately from the stems. Focusing on auditory perception, Post et al. (2008) examined English regular inflections using novel words with (a) potential morphosyntactic information, e.g., *min-minned* /mɪn/-/mɪnd/, (b) with bound morphemes that could be potentially morphological, but not in that specific phonological context, e.g., *rin-rint* /rɪn/-/rɪnt/, and (c) with voiced consonantal affixes that do not normally carry any inflectional or morphosyntactic information, e.g., *plam-plamp* /plæm/-/plæmp/. In this study, the participants took longer to discriminate novel words with potential morphosyntactic information than the novel words violating the rules of English morphophonology; plus, the words with potentially morphosyntactic information took longer than the words with no potential morphosyntactic information (Post et al., 2008). Their results provide evidence that the participants used parsing when processing English inflectional morphology, and they did so with the absence of any lexical meaning of the presented words, i.e., sublexically.[14]

Cilibrasi (2016) built on these findings and created a task that would assess both accuracy and reaction times. Contrary to Post et al. (2008), who used minimal pairs containing bare stems and inflected forms (which is a potentially biased methodology since it may be unintentionally guiding the participants towards morphological decomposition – see Cilibrasi et al., 2019), the novel words in Cilibrasi (2016) had no such limitation, and the task compared only morphologically-inflected forms. The author used only novel words in an auditory decision task, asking the participants to discriminate minimal pairs of inflected novel words. The author attempted to investigate whether stems are stored and processed together with inflectional affixes or separately, and he did so by using novel words containing bound morphemes. Three types of novel verbs were put under analysis: (i) novel words with potential morphosyntactic information in which the inflectional morpheme followed basic rules of English morphophonology (e.g., /veld/ - /velz/), (ii) novel words without potential morphosyntactic information in which the affix could be inflectional but its voicing violated the rules of English morphophonology, i.e., the affix could not be morphological in the given context (e.g., /velt/ - /vels/), and (iii) novel words with a voiced affix that could never be inflectional and carry any morphosyntactic information (e.g., /velb/ – /velm/) – this control condition was added to the experiment to attain for possible voicing effects of the previous two

[14] It is interesting to note that this finding contrasts with work on real existing words. Only the phonological sequences which followed the rules of basic English morphophonology and were therefore allowed in a morphological context (such as *seemed*) were processed more quickly (see, for instance, Marshall & van der Lely, 2006; Korecky-Kröll et al., 2014). Further research should thus focus on the difference between studies with real and novel words.

conditions. The stimuli were then presented to adult native speakers of English. The outcomes of his study showed that participants took longer to discriminate elements with potential morphosyntactic information, similarly to Post et al. (2008). This suggests that monolinguals are sensitive to the presence or absence of morphosyntactic information in items with no semantics. This indicates that morphologically-inflected forms may be decomposed into stems and bound morphemes during perception, even sublexically (Cilibrasi, 2016). Given these findings, the author proposes that the saliency of word-final positions depends on whether any morphosyntactic information is present or not (Cilibrasi, 2016). Similarly to Post et al. (2008), Cilibrasi also found that novel words respecting the rules of English morphophonology took longer to discriminate than novel words violating such rules, which seems to suggest that, when analysing the quality of the affix, monolingual participants also simultaneously analyse the phonological quality of the stem, deciding on its compatibility and consequently on the presence of any morphosyntactic information (Cilibrasi, 2016), as shown for instance in Caramazza et al. (1988).

A subsequent study on bilingual children found a similar pattern in younger learners as well (Cilibrasi & Tsimpli, 2020). In this study, Czech-English bilinguals aged 8 to 11, attending a primary school, were presented with a modified version of Cilibrasi (2016)'s task (the novel words were supposedly uttered by aliens to gauge the participants' attention). The participants were divided into three groups depending on the age of first exposure to English: (i) simultaneous bilinguals exposed to English since birth; (ii) early sequentials exposed to English since nursery; and (iii) late sequentials who were exposed to English only since primary school. The results show that only simultaneous bilinguals displayed the native-like performance described by Cilibrasi (2016), with longer reaction times for morphologically-inflected novel words. The sequential bilinguals, on the contrary, paid attention to the phonotactic constraints of their L1: novel words that were phonotactically legal in both English and Czech were recognised more quickly than the other novel words. This pattern was particularly evident in the early sequentials.

Studies investigating morphological processing at the sublexical level have focused mainly on monolingual speakers, and not much attention has been paid to second-language learners. In 2017, I extended Cilibrasi's study (2016) to more proficient L2 learners of English (with the language level ranging from B1 to C1) with Czech as L1 using the same methodology

(Jiránková, 2017). The study has shown that L2 learners perform similarly to monolingual native speakers – they take longer to process novel words with potential morphological information and discriminate novel words respecting the morphophonological rules of regular English inflection. In both studies, the participants took longer to discriminate novel verbs with potential morphosyntactic than novel verbs without it, suggesting that both English native speakers and more proficient English L2 learners are sensitive to the presence of morphosyntactic information in novel morphologically-inflected verbs and that during perception, these forms might be decomposed, even sublexically (Jiránková, 2017). The study has also shown that the possible decomposition and the potential implicit analysis of the stem are used from a relatively early language level (B1). In my MA thesis, I proposed two possible explanations for these findings: (i) the explicit way of teaching English inflectional morphology in the Czech Republic (where a substantial amount of time is spent on drilling the regular rules of verbal inflection) and (ii) the fact that Czech is inflectionally richer than English and the Czech L2 learners may therefore be more aware of morphology in general when processing other languages (see for instance Ku & Anderson (2003) who show that L1 speakers of a morphologically richer language will be more aware of morphology in general because they are used to paying attention to similar morphological processes in their native language). This strategy might, nevertheless, be "used implicitly by all subjects and is likely a consequence of automatic, unconscious processing" (Jiránková, 2017: 81). Interestingly enough, the study has also reported a significant effect of phonotactic probabilities on reaction times, suggesting that morpheme stripping might not be the only strategy active in the processing of inflectional morphology; some frequency effects might still play a role in the quickness (and potentially even accuracy) of the answer, in line with single-route explanations.

3.2.1.4 Previous L2 studies

Previous studies on second-language acquisition focused primarily on speech production; only recently have researchers begun to delve more into online L2 processing and L2 perception of morphologically complex words. Their findings mostly correlate in showing that native and second-language processing is clearly built on different principles and adult second-language learners' sensitivity to how the word is structured morphologically seem to be less prominent (e.g., Solt et al., 2003; Jiang, 2004; Bliss, 2006; Jiang, 2007; Silva & Clahsen, 2008; Clahsen et al., 2010; Carneiro, 2017), relying more on lexical storage (i.e., on the word as a unit) rather than on morphological decomposition. Experimental studies which report lowered L2 sensitivity to the morphosyntactic information in comparison to native speakers during online language processing involve various methodological approaches, such as self-paced reading (e.g., Jiang, 2004, 2007), timed grammaticality judgements (e.g., Sato, 2007), or eye movements (Keating, 2009). Some of the studies explain this difference as a grammar deficiency in L2 (e.g., Silva & Clahsen, 2008; Clahsen et al., 2010), while others focus more on the influence of L1 features that may be at play (e.g., Solt et al., 2004; Bliss, 2006; Jiang, 2004, 2007). However, some of the online processing studies also report that the native and L2 performance are virtually indistinguishable, with both groups showing the same amount of sensitivity to how the word is structured morphologically and performing decomposition during language perception (e.g., Basnight-Brown et al., 2007; Feldman et al., 2010; Coughlin & Tremblay, 2015; Jiránková, 2017; Coughlin et al., 2019). While some of these studies claim that this sensitivity is not influenced by L2 proficiency and even lower-level participants can perform similarly to native speakers (e.g., Coughlin et al., 2019), some claim that the sensitivity is gained only with higher proficiency (e.g., Ullman, 2005).

The explanation of how L2 learners process morphologically complex words still remains unclear. Many of the studies did not find sufficient evidence for the online decomposition of inflectional morphology in L2 learners (as opposed to native speakers). Ullman (2001, 2005), for instance, distinguishes two different brain memory systems for processing languages – declarative, which involves "the memorization of words in the mental lexicon" (Ullman, 2001: 717), and procedural, "the rule-governed combination of words by the mental grammar" (Ullman, 2001: 717). Building on this distinction, Ullman (2001, 2005) claims that second-language learners depend mostly on the lexical memory (i.e., on full-form storage of inflected words) when processing English inflectional morphology and invoke decomposition less frequently than native speakers (it is either underused or completely absent). These claims find support in a number of other studies, e.g., Silva & Clahsen (2008); Clahsen & Neubauer (2010); or Clashen (2010). On the other hand, a number of previous studies advocate the use of decomposition in L2 online processing (e.g., Hahne et al., 2003; Voga et al., 2014; Kimppa et al., 2019), some of them finding its effects at all proficiency levels (Kimppa et al., 2019), some reporting a slight limitation for L2 learners in comparison to native speakers (Hahne, Müller & Clahsen, 2003). When referring to the two models used when (ir)regular morphologically-inflected words are produced and perceived, there are reports supporting both the dual-route model, which involves the use of both lexical storage and morphological decomposition (Hahne, Müller & Clahsen, 2003), and the single-route model, involving the word's formal and semantic similarity to other words (Feldman et al., 2010).

So far, only a few studies have investigated how sensitive L2 learners are towards inflectional morphemes presented aurally; however, the majority of the findings suggest that the second-language struggle with inflectional morphology also extends to its perception and comprehension (e.g., DeKeyser, 2005). Solt et al. (2004) focused on the perceptual distinction of the regular past-tense allomorphs in L2 learners and native speakers of English. Both L2 groups, beginners and more advanced learners, were able to distinguish the /ed/ allomorph but struggled with the non-syllabic /t/ and /d/. The L2 learners, therefore, did not perceive these allomorphs in the same way as native speakers, showing "perceptual difficulties relating systematically to the phonetic realization of the past tense morpheme" (Solt et al., 2004: 561) even in advanced L2 learners. Jiang (2004, 2007), Peretokina et al. (2014), and Peretokina et al. (2015) report similar difficulties with the perception of word-final *-s* and *-n* morphemes in Chinese L2 learners of English, suggesting that L2 learners can perceive inflectional morphology, though not comparably to native speakers.

Additionally, these findings may be influenced by the learner's L1 and its impact on the L2 processing. Even though L1 transfer in the processing of inflectional morphology has not been systematically investigated so far, recent studies advocate for at least some L1 influence (e.g., Juffs, 1998; White, 2003; Bliss, 2006; Basnight-Brown et al., 2007; Chen et al., 2007; Feldman et al., 2010) and argue that the learner's sensitivity to such morphosyntactic information might be affected by their L1 (see, for instance, Chen et al., 2007). Since Chinese lacks inflectional number and tense marking, the responses of L1 Chinese learners to number or agreement errors may be influenced by the lack of these features in their L1 (Juffs, 1998; White, 2003; Bliss, 2006; Chen et al., 2007; Basnight-Brown et al., 2007). Second-language learners with native languages that are richer in morphology and inflections, in particular, show greater accuracy in grammaticality judgement tasks (Juffs, 1998; Romance languages), tests of verbal morphology (White, 2003; Turkish), or lexical decision tasks (Basnight-Brown et al., 2007; Feldman et al., 2010; both Serbian). While L2 speakers with L1s such as Chinese or Swedish (that are less complex in terms of inflectional mgorphology) may not demonstrate morphological sensitivity, L2 speakers of L1s morphologically richer show morphological facilitation and perform more accurately (see, e.g., Lehtonen & Laine, 2003; Lehtonen et al., 2006; Portin et al., 2006). As Feldman et al. (2010: 121) suggest, "morphological richness in L1 may benefit inflectional morphology in L2." However, some studies also suggest that the lack of a feature in L1 morphology does not necessarily have to pose any significant problems for L2 learners (e.g., Peretokina et al., 2014; Mandarin Chinese) or have a negative effect on the sensitivity to L2 morphological information (Silva & Clahsen, 2008). Peretokina et al. (2015) also state that difficulties in perception do not have to be strictly attributed to the lack of this feature in the participants' L1, but the performance needs to be investigated in connection to phonological and phonotactic factors of both languages. The authors focussed on the detection of the English phoneme /s/ in the coda position in L1 Mandarin speakers since the omission of grammatic inflection is a very common mistake among Mandarin speakers, especially in production. The authors found that in perception, the phonological properties of the coda affected native and L2 participants to a greater extent than its morphological complexity.

On the example of Serbian, Feldman et al. (2010) showed how the complexity of this language may strengthen inflectional processing in English. Similarly to Serbian, Czech is a morphologically rich language, and it is particularly valuable to use this L1 to explore the influence it may have on the mastery of English (L2) inflectional morphology. We can speculate if Czech as a highly inflectional language will produce greater accuracy in the perception of English inflectional morphology. Our findings can be, therefore, directly compared to previous studies operating with less inflectional L1s. Our previous findings (Jiránková, 2017) show that proficient English L2 learners with Czech as L1 are sensitive to the presence of potential morphosyntactic information; in this experiment, we will also look at learners with lower proficiencies. This extension of the previous data, therefore, allows us to look even at the lower-proficiency learners and investigate the potential influence of Czech as a morphologically rich language on English inflectional morphology that is, in comparison, much less complex.

Jiránková's findings (2017), stemming from the data gathered on proficient L2 learners of English, have left several questions unanswered, especially concerning the morphological perception strategies in lower-proficiency Czech L2 learners of English. One of the suggested extensions of the study was to investigate whether decomposition and analysis of the stem during language perception is used from the very beginning of L2 learning or not. This perception experiment aims at addressing this question and thus extends the previous research with data gathered on L2 learners at the A1 and A2 levels, along with the A0-level students who have zero or very little knowledge of English (to possibly account for the presence or absence of potential transfer effects of Czech), and it also attempts to describe the possible changes in how word-final positions are perceived by non-native speakers of English with a highly-inflected L1 that could be possibly still dominant in their performance. The inclusion of a progression of language levels in the experiment is what strengthens our argument in comparison to other L2 studies that usually work with L2 learners as a unit or with two or three selected language proficiencies. Testing participants from the A0 to the C1 level allows us to properly map the development of inflectional morphology in L2 learners and describe its specifics.

3.2.2 Hypotheses

Given previous studies on perceptual processing of inflection morphology, both views of such processing are approached with equal probability. Our study, therefore, operates with several possible findings: (a) if morphologically-inflected words are stored as units with the stem and the bound morpheme united as one, and they are elicited as such during perception (see the atomist view above), the quickness of the reaction will be determined by the phonotactic probabilities of individual stimuli and items with high phonotactic probabilities will take less time to be discriminated; (b) if inflections and stems are stored separately and decomposed into two separate parts during perception (see the decompositional view above), morphosyntactic information will determine the quickness of the reaction – the participants will take longer to discriminate novel words with potential morphosyntactic information (Cilibrasi, 2016; Jiránková, 2017; Cilibrasi et al., 2019).

Given the previous findings by Jiránková (2017), two possible outcomes may arise for the lower-proficiency levels: (a) even lower-proficiency Czech L2 learners may be more aware of morphology in general since their L1 is morphologically rich and the sensitivity to morphological processes in their L1 is essential (see Ku & Anderson, 2003), and they may thus decompose morphologically-inflected forms during perception, or (b) lower-proficiency Czech

L2 learners may not decompose morphologically-inflected forms since storing the whole form as a unit may be more economical at this stage of L2 acquisition, or they may not be simply proficient enough to know the English inflectional morphemes and be able to distinguish it aurally in spoken language.

Building on previous studies on sublexical perception (e.g., Post et al., 2008; Cilibrasi, 2016; Jiránková, 2017; Cilibrasi et al., 2019), we predict that there will be a difference in the quickness of reaction between novel words with and without potential morphosyntactic information and that the participants will take longer to distinguish the novel words with the morphosyntactic information present. Given previous findings by Jiránková (2017), we also expect the possibility of frequency effects influencing the reaction times. Based on Cilibrasi et al. (2019), we also expect that two identical items in a minimal pair will be easier to judge than the slightly different items (see also Beauvillain, 1994; or McQueen & Cutler, 1998).

3.3 Methodology

The methodology, the set of stimuli, and the general procedure of the task is taken from Cilibrasi (2016) and Jiránková (2017) to ensure comparable findings.

3.3.1 Ethics

The experiment was approved by the Charles University Ethics Committee in 2016 when the permission to proceed was given to the initial experiment conducted on more proficient L2 learners (see Jiránková, 2017).

3.3.2 Recruitment

The participants were recruited via an e-mail. Most of the second-language learners were the students at the Faculty language school since the language school had recently tested these participants for their language level in English by a placement test.[15] The participants with a required language level and first language were addressed directly. A small fee was given to all participants (as this research project has been financially supported by a Specific Academic Research grant, "Jazyk a nástroje pro jeho zkoumání").

The C1-level students were recruited via a short Facebook ad and tested with the same placement test. Three B1-level students were recruited at a secondary school in Kolín and tested in the same way for their language level. The use of dictionaries and other study materials was prohibited during the testing.

3.3.3 Participants

91 English second-language learners with Czech as their first language at A1 - C1 proficiency levels (16 at the A1 level, 15 at the A2 level, 20 at the B1 level, 20 at the B2 level, and 20 at the C1 level) and 4 A0-level learners with zero or very little knowledge of English were recruited. None of the participants had been diagnosed with any language impairment or hearing deficits. The mean age was 34.6 years, standard deviation then 10.1 years. 75 were female, 20 were male.

3.3.4 Consent

The testing took place in a quiet room located at the Faculty of Arts or at a secondary school in Kolín. The participants were asked to express their consent with the experiment, data anonymization, and research goals. They were also informed that all the data would be stored in a locked cabinet to which only the administrators would have access. Each participant was then assigned a random five-digit number[16] in order to anonymize their data.

---

[15] The placement test was created by Dr Tomáš Gráf for levels A1 to C1 and was based on the data available in the English Grammar and Vocabulary Profiles. These two databases describe the average grammatical and lexical knowledge of an L2 learner at a given language level (as divided by the Common European Framework of Reference). The placement test is available only to the teachers and members of the Faculty Language school through https://dl1.cuni.cz/course/view.php?id=424#section-1.

[16] Generated at https://www.random.org/.

3.3.5 Stimuli

The study adopted 120 novel verbs from Cilibrasi (2016), along with the values of positional segment frequency and the biphone frequency (for further description, see below).

3.3.5.1 The creation of the stimuli

Following the above-mentioned rule for regular verbal inflection (with a voiceless bound phoneme following a voiceless stem-final phoneme and a voiced bound phoneme following a voiced stem-final phoneme), it was possible for Cilibrasi (2016) to create (i) novel verbs with bound inflectional morphemes that followed the rules of regular English morphophonology (e.g., /vɪlz/ or /naɪld/) and (ii) novel verbs with bound inflectional morphemes that could be potentially morphological, but in a different phonological context (e.g., /vɪls/ or /naɪlt/) – here, a voiced phoneme would have to follow to create a potentially inflectional bound morpheme.[17] Since this experiment investigates possible morpheme stripping effects at the sublexical level, attention was paid to the discrimination of those two groups of novel verbs. The phonotactic legality of all novel verbs designed by Cilibrasi was checked using a Phonotactic Probability Calculator (Vitevitch & Luce, 2004). This software calculates both phonotactic probability values for given phonological units (so-called *positional segment frequency*, PSF, i.e., how phonotactically probable it is that a particular phoneme will appear in the given word position) and segment sequences (so-called *biphone frequency*, BF, i.e., how probably it is that two adjacent sounds will appear together in a word) from Webster's Pocket Dictionary (Anderson & Byrd, 2008) – for the PSF and BF values of the presented stimuli, see Appendix 1. Having imported the novel words into the calculator, Cilibrasi obtained the values of positional segment frequency and biphone frequency for all his novel words.

The stimuli start with four different consonants: /v/, /n/, /θ/, and /dʒ/ - these phonemes are phonotactically legal at word beginnings in English, while their frequency is still rather low to reduce the risk of the novel verbs already existing in English (Cilibrasi, 2016). This was useful even more so that all the novel verbs are monosyllabic. Each novel word contains only

[17] One possible counterargument to this is that verbs such as *felt* /felt/ still carry morphological information even though the voiceless affix follows a voiced consonant of the stem. These verbs are, however, considered irregular and marked and exceed the scope of this thesis.

one vowel among the following: /ɪ/, /aɪ/, /æ/, /ɔ/, /ʌ/ - all of them phonotactically legal in word-second positions. The stimuli then end in (i) /lz/ and /ld/ codas that follow the rules of regular English morphophonology and can be morphological and (ii) /ls/ and /lt/ codas that do not follow the rules of regular English morphophonology and are thus potentially non-morphological. The monosyllabicity of the novel words was pertained on purpose to reduce the risk of a further morphological decomposition during the task (Cilibrasi, 2016).

Combining each onset (4 consonants) with each nucleus (5 vowels) then produced 20 novel-word bases. The four codas were then added, creating 80 novel verbs in total. As described above, half of them contained potential morphological information (those ending in /lz/ and /ld/) and a half of them did not (those ending in /ls/ and /lt/). This created two different conditions where the contrast between the potential morphological and non-morphological minimal pair can be explained by the quality (voiced or unvoiced) of the final bound morpheme in connection to the preceding voiced consonant (/l/). In order to control for the fact that the difference between the two conditions could be purely phonological (based on the difference in voicing), a third condition was created. Here, the final phonemes are both voiced as in the first condition, but the bound morpheme cannot be morphological under any circumstances (even in any other context – as is the case in the second condition): /lb/ and /lm/. Adding these two voicing-control codas to the initial bases generated another 40 novel verbs. In total, the stimuli dataset contained 120 novel verbs. An overview of all conditions is summarized in the table below (for the full list of stimuli, see Appendix 1):

Condition	**Potentially morphosyntactic minimal pairs**	**Non-morphosyntactic minimal pairs**	**Voicing control condition**
Example	/vɪld/ - /vɪlz/	/vɪlt/ - /vɪls/	/vɪlb/ - /vɪlm/
Features	• final bound morpheme voiced, following a voiced consonant • → morphological ending	• final bound morpheme voiceless, following a voiced consonant • → non-morphological ending that could be morphological in a different context	• final bound morpheme voiced, following a voiced consonant • → non-morphological ending

Table 1. Types of minimal pairs adapted from Cilibrasi (2016) divided into conditions (sourced from Jiránková, 2017).

The stimuli were presented as minimal pairs – "pairs of words in which a difference in meaning depends on the difference of just one phoneme" (Roach, 2000: 63) – exemplified for instance by the difference between *kin* /kɪn/ and *king* /kɪŋ/ – more specifically, as morphosyntactic minimal pairs (Law & Strange, 2010), in which the items "differ in one phoneme that is also an inflectional morpheme" (Cilibrasi, 2016: 49). The use of minimal pairs also suits well the study of perception (Cilibrasi, 2016). The dataset was therefore further divided into 60 minimal pairs consisting of different items (the items differed in the final bound morpheme) and 60 minimal pairs of identical items – 20 pairs for each condition described above in each type.

Regarding the pairing of individual stimuli items, the novel words of the "same" type always ended in a plosive consonant (identified by odd numbers in Appendix 1), while in the "different" type, the first novel words ended in a plosive consonant, while the second one ended in its corresponding nasal or fricative (identified by an odd number and a following even number in Appendix 1). The second novel word in each minimal pair never ended in a plosive to ensure that the first novel word would always prime "some sort of *verb-like activation* in the morphological condition" (Cilibrasi et al., 2019: 755) (since the novel verbs ending in plosives can potentially trigger the third-person-singular or plural-noun priming). The same pattern was preserved in all three conditions (Cilibrasi et al., 2019). Since the dataset consists of 60 novel words with identical items and 60 corresponding novel words with different items, the participants were, in consequence, presented with 120 trials in total.

It is essential for further analysis to mention that mean durations (in ms (SD)) for each condition vary: 224 (45) for the morphological condition, 146 (37) for the non-morphological condition, and 178 (33) for the phonological control. This variation in duration is unavoidable if we want to create realistic English stimuli since morphologically-inflected verbs undergo vowel lengthening in English if the stem-final vowel is voiced (which is the case in all our stimuli). There lies a potential confound of condition and stimulus length since the morphological condition is also the longest (and potentially longer reaction times to this condition may be explained both by the presence of morphosyntactic information and the duration of the stimuli). However, this effect is mitigated by the inclusion of both stimulus length and condition as independent variables in the model used for the analysis.

3.3.5.2 Stimuli recording

The stimuli were recorded in the sound booth of the School of Psychology and Clinical Language Sciences at the University of Reading by a native speaker of English who had been previously instructed to produce vowel lengthening preceding a lenis and pre-fortis shortening in the morphological condition (Cilibrasi, 2016).

3.3.6 Procedure

During the experiment, each participant was presented with 120 minimal pairs. They were told they would hear two words and that they should decide whether the words differ or not by pressing one of the assigned keys.

The experiment was created in the psycholinguistic software PsychoPy v1.84.0 (Peirce et al., 2019). The experiment took place in a quiet room, with participants wearing headphones and following the instructions on the computer screen. To avoid confusion in participants with lower proficiency, the instructions were presented in Czech only.[18] The participants were told they would hear two words and asked to press one of the assigned keys as soon as possible if they think the two words are the same/different. Each participant was first presented with one practice trial with one aurally-presented minimal pair to get accustomed to the task.

The trial session was followed by a test session, in which the participants were presented with a decision task (same vs. different). In total, the participants heard 120 trials (minimal pairs) of novel verbs with three conditions. In each trial, the participant was aurally presented with a minimal pair consisting either of identical or different novel-word items. All minimal pairs were presented randomly, and the order of stimuli was different for each participant. The

[18] During the instruction phase, the participants were not made aware of the goal of the experiment, neither the language it was supposed to test. From the recruitment, they only knew they needed to be at a particular language level to be included in the testing. Since the allomorphic rules for English inflectional morphology are usually not taught explicitly in the Czech Republic and we did not expect the participants to have any explicit knowledge of it, our aim was to focus on the implicit knowledge of the rules and the participants' sensitivity to the presence of morphology. In Czech, the voicing assimilation rules are mostly regressive (or anticipatory) when the assimilation is governed by the last phoneme in the group (e.g., vztek [fstek] where the voiceless *t* causes the devoicing of the preceding phonemes). However, instances of progressive assimilation can be found in Czech as well – for instance, tři [tř̥i], in which the voiceless *t* causes *ř* to become devoiced as well (Internetová jazyková příručka).

recorded measures obtained from this experiment included the reaction times to the presented novel verbs for a reaction-times analysis and pressed keys for an accuracy analysis.

3.4 Results

3.4.1 Reaction-times analysis

The reaction-times analysis is designed to answer the question of whether reaction times, i.e., the length of time taken by a participant to respond to a given stimulus, are different for different minimal pairs – either because of their type (same vs. different) or condition (potentially morphological, non-morphological, and control condition). These measures offer us a close look into possible storage and perceptual retrieval of English inflectional morphology, helping us adjudicate between the decompositional and atomist approach.

3.4.1.1 Editing the dataset

A total of 10,986 reaction times were recorded (448 for the A0 level; 1,825 for the A1 level; 1,709 for the A2 level; 2,302 for the B1 level; 2,366 for the B2 level; and 2,336 for the C1 level). The trials were first checked for outliers, i.e., observations "that lie outside the overall pattern of a distribution" (Moore & McCabe, 1999) and "which fall more than 1.5 times the interquartile range above the third quartile or below the first quartile" (Ambrus et al., 2017: 3). For the presented data, the first quartile (FQ) starts at 1.13 seconds, and the third quartile (TQ) ends at 1.43 seconds, thus creating an interquartile range (IQR) of 0.3 seconds. Using the definition mentioned above of an outlier, the upper-bound for the regular data equals 1.5*(TQ + IQR), i.e., 1.88 seconds. The lower-bound then equals 1.5*(FQ – IQR), i.e., 0.68 seconds. All data points that did not fit into the range of 0.68 and 1.88 seconds were then considered outliers and deleted from the final data set. Since our primary interest was the speed of reaction with a correct response to a given stimuli pair, the reaction times for incorrect responses were excluded from the dataset as well. After the omission of the outliers and incorrect responses, the dataset had the following characteristics:

	N	Mean reaction times (s)	Minimum reaction time (s)	Maximum reaction time (s)	Standard deviation (s)	Mean reaction times for condition 1 (morpho) (s)	Mean reaction times for condition 2 (non-morpho) (s)	Mean reaction times for condition 3 (control) (s)
A0 level	392	1.293	0.75	1.86	0.217	1.359	1.207	1.322
A1 level	1,741	1.390	0.87	1.88	0.199	1.464	1.287	1.424
A2 level	1,648	1.343	0.69	1.87	0.205	1.417	1.245	1.371
B1 level	2,217	1.268	0.71	1.88	0.202	1.344	1.177	1.285
B2 level	2,296	1.226	0.69	1.88	0.199	1.313	1.131	1.234
C1 level	2,271	1.233	0.69	1.88	0.216	1.309	1.146	1.248
All part.	10,565	1.284	0.69	1.88	0.214	1.368	1.199	1.314

Table 2. Descriptive statistics of the dataset used for the reaction-times analysis.

For each trial, reaction times and pressed keys were recorded (along with the ID number, the language level, the placement test score, the type (same vs. different), the condition (morphological, non-morphological, control), the length of the minimal pair (in seconds), and the positional segment and biphone frequencies).

Given the size of the analysed sample that consists of 95 participants, the analysis assumes normal distribution based on the central limit theorem, which states that "the means of a random sample of size, *n*, from a population with mean, *μ*, and variance, *σ2*, distribute normally with mean, *μ*, and variance, *σ2/n*" (Kwak & Kim, 2017: 144). The more the sample size increases, the closer its mean is towards the population mean. The central limit theorem then applies whenever the sample is large enough – generally, $n \geq 30$ – and the sample distribution is then considered normal (Kwak & Kim, 2017). This holds true even if the population distribution was skewed. In the case of a sufficiently large data sample, the analysis can work with parametric tests, which "produce more accurate and precise estimates with higher statistical powers" (Kwak & Kim, 2017: 144) than non-parametric tests.

To investigate whether the data showed significant effects of various independent variables on reaction times, the results were analysed with a linear mixed-effects model in R (R

Development Core Team, 2011)[19]. The linear mixed-effects model was chosen to investigate manipulated variables and random effects that may appear as a result of individual differences among participants since the model works with individual trials rather than average scores. The model also effectively deals with missing data and atypical data distribution and can treat participants and individual stimulus items as random effects (Baayen et al., 2008; Blything et al., 2018). The dependent variable used was the reaction times to presented stimuli; the independent variables (fixed factors) included (a) the level, (b) the condition (morphological, non-morphological, control), (c) the type (minimal pairs with identical or different items), (d) the length (duration) of the minimal pair (measured in seconds), and (e) the positional segment frequency and biphone frequency. Independent variables were imported into the model simultaneously to investigate their potential interaction. To account for possible multicollinearity of the independent variables, the variance inflation factor (VIF) was checked for each model. The values for the decompositional analysis were as follows: condition - 1.103; type - 1.226; length -1.332; level - 1.001; and for the atomist analysis as follows: positional segment frequency - 1; biphone frequency - 1; level - 1.[20] The model accounted for random effects by keeping their structure in its simplest and most-frequently-used form of (1|part) + (1|item) with "part" standing for participants and "item" for individual novel words. Only reaction times with correct answers were used in the analysis.

3.4.1.2 The decompositional analysis

As previously outlined in our hypotheses, if stems and inflections are stored separately, and morpheme stripping takes place during perception (see for instance Berko, 1958; Pinker & Ullman, 2002; Guasti, 2004), the potential morphosyntactic information will be the main predictor of how much time the participants will need to discriminate the pairs of novel words: the presence of the morphological information will result in longer reaction times.

In the initial analysis, we followed Pérez et al. (2016)'s procedure of obtaining the most precise and explanatory models. Using the random effects for participant and item (see (1|part)

[19] The data meet all the necessary assumptions of the model.

[20] VIF values are usually ≥ 1 (signalling the absence of multicollinearity), and while there is no exact limit to multicollinearity values, VIF values exceeding 5 to 10 are often an indication of a problematic collinearity and the variables in questions should be removed since their information is redundant in combination with other variables (James et al., 2013).

+ (1|item) above), we started the analysis with the highest number of possibly interacting independent variables: level, condition, type, and length of the stimuli, and compared it with simpler interactions using a stepwise model comparison (see Pérez et al., 2016). Since the deletion of this complex four-way interaction did not give a statistically significant outcome ($p = 0.541$), it was removed from the model. Following the same procedure, other interactions were removed from the model as well: the interaction among condition, type, and length ($p = 0.27$); condition, type, and level ($p = 0.952$); type, level, and length ($p = 0.136$); condition and type ($p = 0.171$); type and length ($p = 0.446$); length and level ($p = 0.31$); and type and level ($p = 0.115$), keeping only the interaction between the condition and the level ($p = 0.001$ **) and between the condition and the length ($p = 0.02$ *). Since the removal of any of the two interactions and of the individual variables from the model resulted in significant results of the ANOVA ($p = 0.026$ *), the final model included the condition, level, length, and type as independent variables and participants and items (i.e., the novel words) as random effects in the following R code: cond:length + cond:level + cond + type + length + level + (1|part) + (1|item). P-values were generated with the *lmerTest* package (Kuznetsova et al., 2017).

First, an analysis was run on the whole dataset to see if participants at different levels performed differently and whether there was any effect of the level, condition, type, or novel-verb length (and therefore, if individual post-hoc tests were needed). The overview of the linear model can be seen in Appendix 1. To get a simpler overview of the significant variables, an analysis of variance (ANOVA) was then run on the model. The analysis showed a significant p-value for all independent variables and for the interactions between condition and length and condition and level (for detailed results, see the table below):

	Df	Sum Sq	Mean Sq	F value	P-value
Condition	2	0.203	0.101	4.134	**0.02** *
Type	1	0.162	0.162	6.627	**0.01** *
Length	1	2.190	2.190	89.351	**< 0.001** ***
Level	5	0.832	0.167	6.792	**< 0.001** ***
Condition:length	2	0.339	0.170	6.923	**0.001** **
Condition:level	10	0.519	0.052	2.118	**0.02** *

Table 3. The analysis of variance (ANOVA) table for the decompositional analysis. The numbers in bold indicate a significant effect. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

These results suggest that participants at different language levels reacted differently to the presented stimuli, which hints at some form of language development. The findings also suggest that the participants reacted differently to the three conditions, which hints at the possible effect of the morphological information on reaction times. The reaction times were also influenced by the length of the presented stimuli pair, which suggests that duration was a decisive factor in the performance. The significant effect of the stimuli type on reaction times indicates that the participants reacted differently to the stimuli with identical and different items. The significance of the interaction between the condition and the level suggests that participants at different language levels react differently to different conditions, and the significance of the interaction between the condition and the length suggests that reaction times to different conditions were also influenced differently by the novel verb's length. This provides us with enough evidence to perform post-hoc tests and see how the results differed at each condition, while the rest of the findings asks for graphical visualization to make us understand the relation between the variables.

3.4.1.2.1 The role of the language level

The initial ANOVA has shown a significant effect of the language level, which suggests that participants at different levels reacted differently to the presented stimuli. The figure below visualizes the participants' performance at different language levels. The plot shows a gradual decline in mean reaction times with higher proficiency, with the exception of the A0 level (at which the participants reacted quicker than at the A1 and A2 levels) and the B2 level (at which the participants performed quicker than at the C1 level).

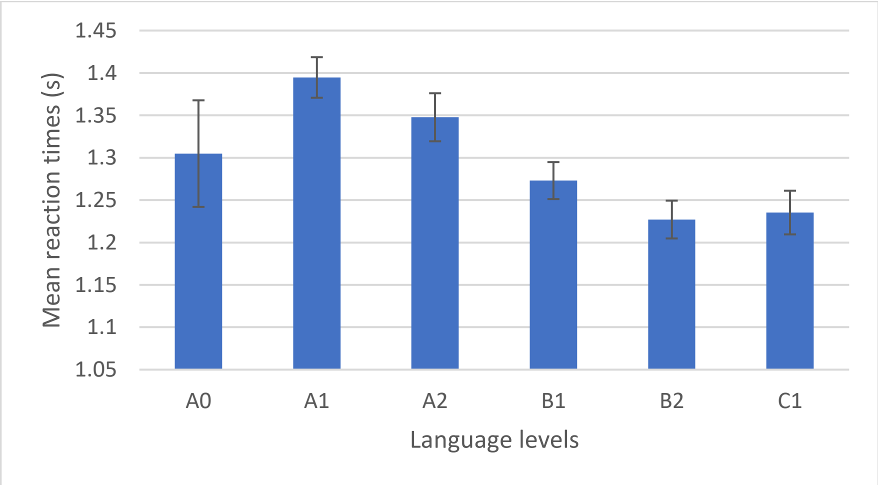

Figure 1. Mean reaction times at each language level.

To better understand the relationship among the language levels, independent samples t-tests were run between individual language levels to see if their difference turns out significant or not. The summary below shows that the A0 level performed in a similar way to all other language levels; however, this result might be highly influenced by the low number of participants involved (see also the discussion below and the drawbacks section of the Conclusion). The A1 and A2 levels then performed in a significantly different way from the rest of the language levels, but not from each other (the lowest language levels thus seem to be performing similarly to one another). The B1, B2, and C1 levels performed similarly to one another. The table below shows an overview of the post-hoc t-test run among individual levels:

	A1	A2	B1	B2	C1
A0	t (119) = -1.33, p = .253, two-tailed	t (119) = -0.62, p = .568, two-tailed	t (119) = 0.48, p = .66, two-tailed	t (119) = 1.17, p = .308, two-tailed	t (119) = 1.02, p = .364, two-tailed
A1		t (119) = 1.26, p = .216, two-tailed	t (119) = 3.756, **p < .001**, two-tailed	t (119) = 5.12, **p < .001**, two-tailed	t (119) = 4.53, **p < .001**, two-tailed
A2			t (119) = 2.09, **p = .046**, two-tailed	t (119) = 3.347, **p = .002**, two-tailed	t (119) = 2.93, **p = .006**, two-tailed
B1				t (119) = 1.47, p = .149, two-tailed	t (119) = 1.12, p = .271, two-tailed
B2					t (119) = -0.242, p = .810, two-tailed

Table 4. Results of the post-hoc t-tests run among individual language levels. Significant results are indicated in bold.

This result is, indeed, interesting since it seems to suggest that proficiency in L2 generates some effect on reaction times only between the lower and higher proficiency levels – the lowest language levels (A1 and A2) performed differently from the more proficient language groups (B1 to C1), while the higher-proficiency participants showed no significant difference in performance. However, looking back at the plot above, it should be noted that the B1 level might not show any significant difference from the B2 and C1 levels statistically; yet the B1 participants still reacted slower to the presented stimuli. The A1 and A2 levels then also perform similarly to one another but differently from the B1, B2, and C1 levels. The figure above shows that their reaction times are significantly longer than those of the more proficient levels. The language level thus seems to play a role in the quickness of reaction to the presented stimuli, and some knowledge of English is probably needed to attain a quicker reaction.

3.4.1.2.2 The role of type

The initial ANOVA has also shown a significant difference between the participants' reaction times to minimal pairs with identical and different items. The pairwise comparison of reaction times to the two stimulus types has shown a significant difference between the reaction times in each stimuli type ($t(59) = 2$, $p = .02$, two-tailed), with the different type taking longer to be discriminated (based on the figure below and on the sign of the test statistic). The figure below depicts a graphical visualization of the performance.

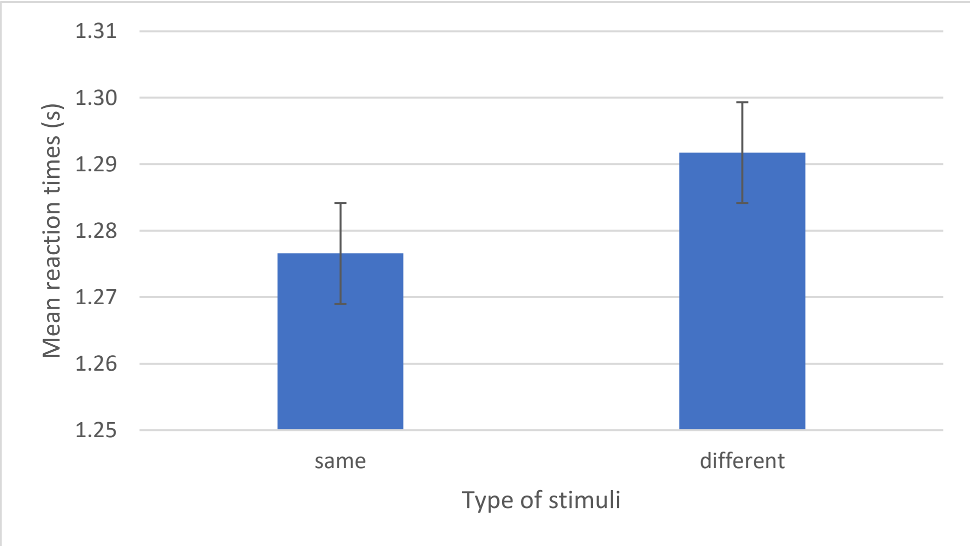

Figure 2. Mean reaction times to the type of the presented stimuli.

The plot clearly shows that the participants reacted quicker when presented with the minimal pair consisting of identical items, while the reaction to the minimal pair with different items took longer. The longer reaction to those items could be explained by the general fact that recognising a small difference in two similar items is more costly than recognising two items as being identical, as shown in previous research (McQueen & Cutler, 1998; Cilibrasi et al., 2019). The phonological difference between the items in the stimuli set was also placed at the very final position in the novel word and was therefore more cognitively demanding to process.

3.4.1.2.3 The role of the condition

The initial ANOVA has also shown a significant effect of the condition on reaction times, suggesting that participants react differently to different conditions. The interaction among the conditions was first visually depicted in the figure below.

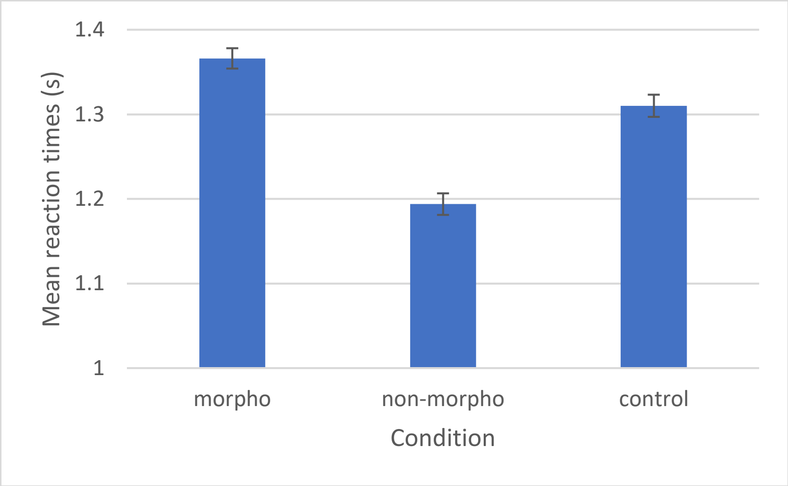

Figure 3. Mean reaction times to the three conditions (morphological, non-morphological, and control conditions).

The plot shows that out of the three conditions, the potentially morphological condition takes the longest to be distinguished, followed closely by the control condition and lastly by the non-morphological condition. This visualization suggests that reacting to the potentially morphological stimuli was more demanding than reacting to stimuli without encoded morphological information. This finding hints at the participants' sensitivity to the presence and absence of morphology and also at the fact that the participants might simultaneously analyse both the phonological quality of the bound affix and of the final position in the stem (since the reactions to the morphological and control conditions, which respect morphophonological rules of English inflection, took longer to discriminate than the non-morphological condition, which ultimately violates these rules).

Independent samples post-hoc t-tests were run among individual conditions to see whether their differences turn out significant or not and to make us understand their relation. The results (summarized in the table below) show that each condition is significantly different from the other two conditions. Since reaction times to the three conditions are significantly different from one another (and especially the morphological condition from the non-morphological condition), morpheme stripping (or at least morpheme implicit recognition) might have been taking place during the perceptual processing of the stimuli.

	cond 2 (non-morpho)	cond 3 (control)
cond 1 (morpho)	t (39) = 9.83, **p < .001**, two-tailed	t (39) = 3.15, **p = .002**, two-tailed
cond 2 (non-morpho)		t (39) = -6.36, **p < .001**, two-tailed

Table 5. Results of the post-hoc t-tests among individual conditions. Significant results are indicated in bold.

3.4.1.2.4 The role of length

The length (or duration) of the presented stimuli has also shown a significant effect on the reaction times in the initial ANOVA. The figure below visualizes the relation between reaction times and the length (duration) of the stimuli.

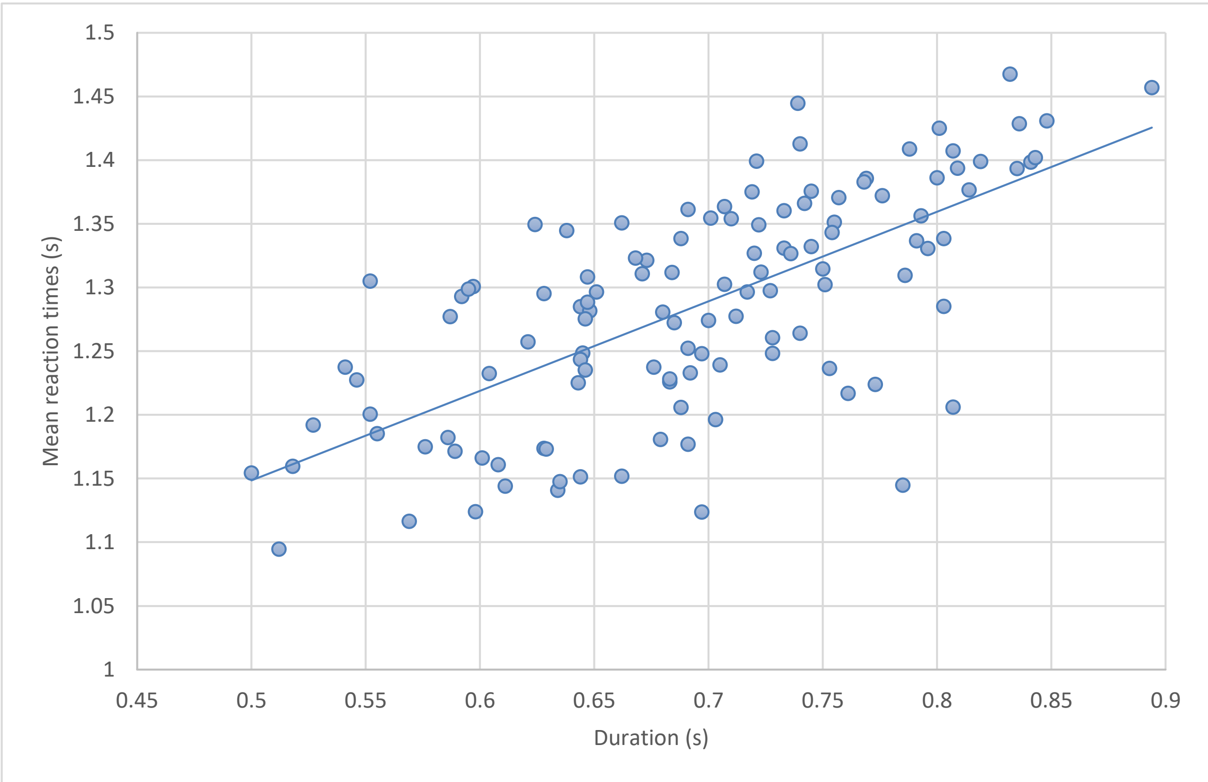

Figure 4. Mean reaction times as a function of the length (duration) of the presented stimuli.

The plot shows that the shorter the novel word is, the quicker the reaction; on the contrary, with longer novel words, the reaction takes longer. The longer stimuli thus seem to be cognitively more demanding, and the pre-lenis lengthening seems to influence the reaction to the presented stimuli. This finding is rather interesting since it shows that the length of the novel word as a unit influences the quickness of the reaction, and that the duration of the stimuli is a decisive factor in the performance. In connection to the previous results on the effect of condition, these new findings suggest that the performance is dependent on the absence or presence of bound morphology, but also on duration, with the performance being influenced by the stimuli length. Its interaction, therefore, needed to be explored in greater detail in further analysis.

3.4.1.2.5 The interaction between condition and length

The next significant interaction in the initial ANOVA was found between the condition and the length (duration) of the stimuli, suggesting that participants reacted differently to novel verbs of different lengths in different conditions. A visual representation of the interaction was, therefore, needed to understand the difference among the conditions in connection to the stimuli duration (this can be seen in the figure below):

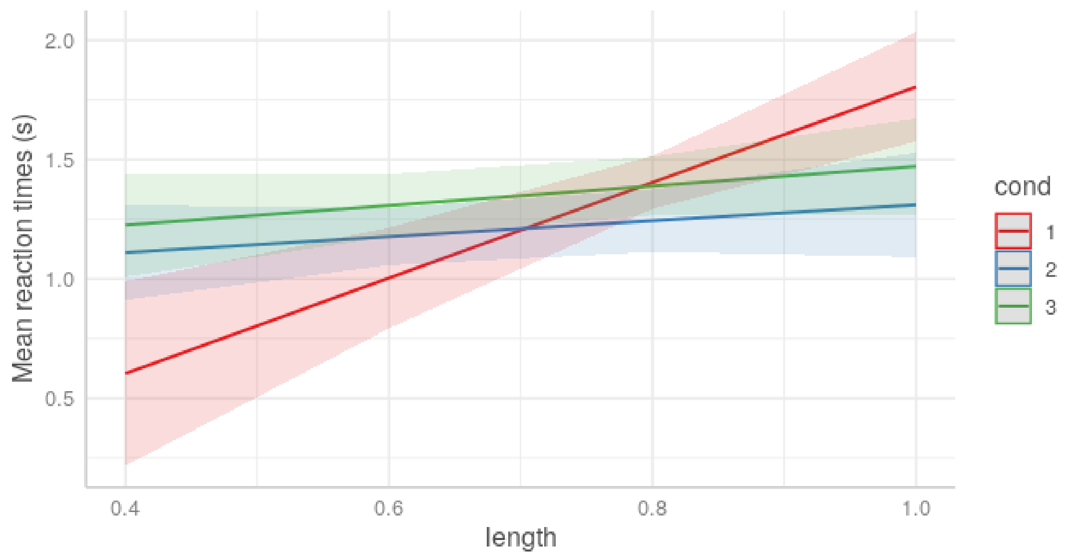

Figure 5. Mean reaction times divided by the condition as a function of the stimuli duration (cond 1 = morphological, cond 2 = non-morphological, cond 3 = phonological control condition).

The plot above shows that for all conditions, the quickness of reaction decreases with the length of the stimuli. Out of the three conditions, morphological condition 1 is the condition that is the most sensitive towards stimuli duration (based on the apparent steepness of the trendline). A deeper analysis is, however, needed to understand the interaction (and the differences between the conditions) fully. The dataset was therefore divided into two halves: (i) stimuli with shorter and (ii) with longer duration and it was plotted for a proper understanding of the situation:

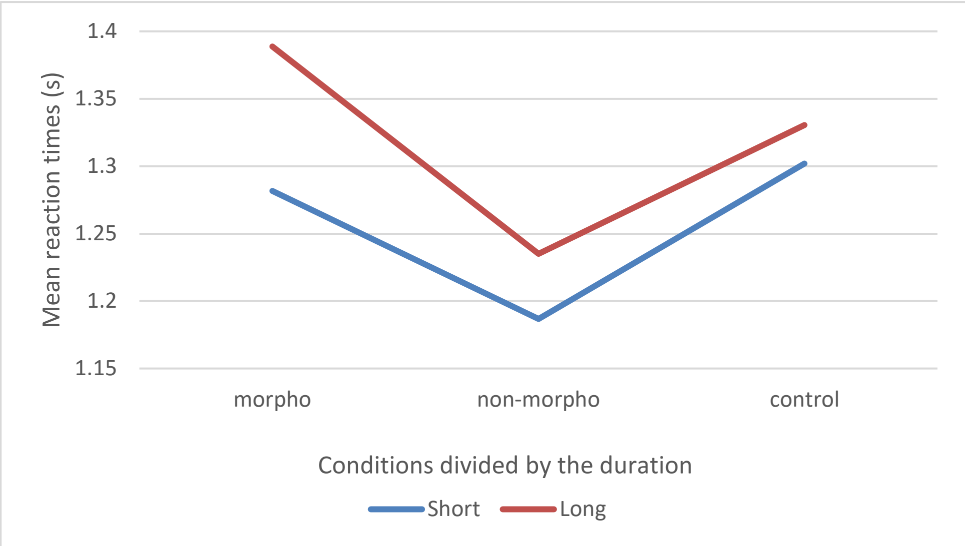

Figure 6. Mean reaction times to each condition divided by the duration of the stimuli.

The plot above shows that for novel verbs with shorter duration, the control condition takes the longest to be discriminated, followed by the morphological and the non-morphological condition. For the longer-duration novel verbs, it is the morphological condition that takes the longest to be discriminated, followed by the control and the non-morphological condition. For both types of duration, the non-morphological condition thus takes the shortest time to be discriminated, but the two subsets differ in the quickness of reaction to condition 1 (morphological condition) and condition 3 (control condition). The morphological condition is slower only with long novel words, not with short ones. This indicates that the performance in response to the shorter-duration stimuli is more influenced by the duration and voicing, with the control condition showing slower reaction times than the morphological and non-morphological one, while with the longer-duration novel words, the reaction times to the morphological condition are considerably slower than for the two remaining conditions. This finding (the evident slowness of the morphological condition with stimuli of longer duration) could potentially somewhat lessen the confound of the condition and length effect on the reaction time (though not fully) since it shows that even if the stimuli were of longer length, the reaction to the stimuli in the morphological condition was still taking the longest to be discriminated.

3.4.1.2.6 The interaction between the condition and the level

The last significant interaction of independent variables in the initial ANOVA was the interaction between the condition and the language level, which suggests that participants at different language levels reacted differently to different conditions.

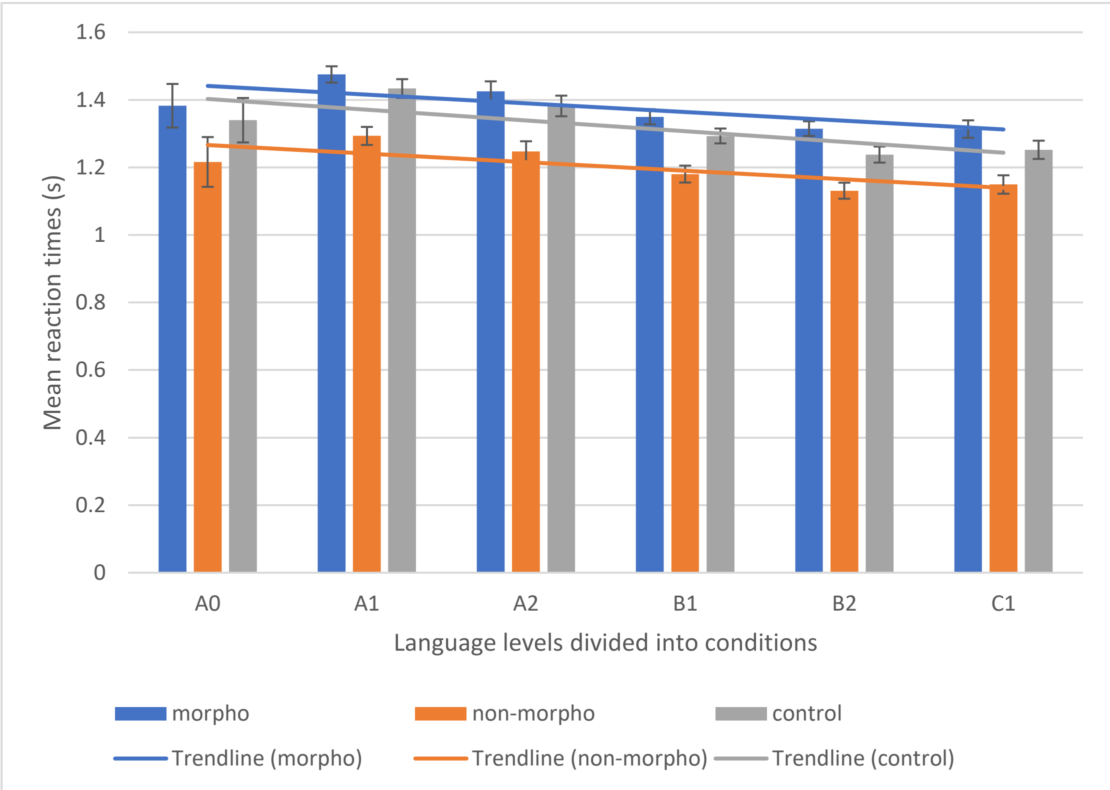

Figure 7. Mean reaction times to the three conditions divided by the language level.

The graphical visualization of the data did not show any apparent differences among the language levels and their performance in each condition – all subjects still preserved the same condition pattern: the morphological condition taking the longest to be distinguished, closely followed by the control and the non-morphological condition. The reaction times to the morphological condition also keep a lowering tendency with higher proficiency (with the exception of the A0 level), hinting at the gradual speeding of reaction to the presence of morphology in the course of second-language acquisition. The trendline of the morphological condition also shows the highest mean reaction times in comparison to the other conditions, suggesting that the morphological condition truly is the most demanding for all proficiency levels. The progression in the two remaining conditions remains in decline, with the non-morphological condition producing the lowest mean reaction times, followed by the control condition. This progression is kept for each language level.

To understand the interaction between level and condition in greater detail, a data subset was created for each condition at each language level, and ANOVAs were run among conditions at individual proficiency levels. The results can be seen in the table below:

Language level	ANOVA results
A0 level	$F(2, 9) = 1.61$, $p = .253$
A1 level	$F(2, 45) = 13.20$, $\mathbf{p < .001}$
A2 level	$F(2, 42) = 9.48$, $\mathbf{p < .001}$
B1 level	$F(2, 57) = 13.91$, $\mathbf{p < .001}$
B2 level	$F(2, 57) = 15.94$, $\mathbf{p < .001}$
C1 level	$F(2, 57) = 9.63$, $\mathbf{p < .001}$

Table 6. Results of the ANOVAs run among individual conditions at each language level. Significant results are indicated in bold.

The results indicate that the difference in performance among the language levels comes down to the A0 level reacting similarly to the three conditions, while the rest of the proficiency levels keeps the pattern outlined above in section 3.4.1.2.3. This might be, again, explained by the low number of subjects participating in the experiment and their individual differences. However, we may speculate whether some (at least basic) knowledge of English is not needed for the proper perceptual distinction of regular English word-final allomorphs and eventually for distinguishing the presence or absence of bound morphosyntactic information. The morphological richness of Czech thus might not be enough to ensure suitable perceptual performance.[21] It is, nonetheless, interesting to observe the overall pattern of the reaction times to individual conditions and their spread over the language levels as it is kept relatively constant among the proficiency levels and a similar pattern of reaction is observed at the lowest A1 level and the proficient C1 level, but also in the native speakers tested by Cilibrasi (2016). This might hint at a possible facilitatory effect of Czech and its rich morphological endings on the morphological awareness of Czech L2 learners even at the lowest proficiency levels.

3.4.1.3 The atomist (unit-storing) analysis

The second possible hypothesis of inflection processing operates with the notion that if stems and bound inflectional morphemes are stored together as units (e.g., Tomasello, 2000), the phonotactic probabilities of individuals stimulus items will be the decisive factors in how much time the participants will need to distinguish the novel words. Shorter reaction times are also expected for the novel words with higher phonotactic probabilities.

Once again, the data were analysed with a linear mixed-effects model. The dependent variable used was the reaction times to presented stimuli; the independent variables (fixed factors) included (a) the level, (b) the positional segment frequency (FPS) / the biphone frequency (BF). First, an analysis was run with the interaction of level and PSF and level and BF to see if the performance was influenced by the stimuli frequencies and whether participants at different levels performed differently when presented to the stimuli with different frequencies. The overview of the linear mixed-effects models for both PSF and BF values can be seen in Appendix 1. The analysis of variance (ANOVA) showed no significant p-value for the effect of the level and BF on reaction time (level: $p = 0.982$, BF: $p = 0.443$, level:BF: $p = 0.989$), but it showed a significant p-value for the effect of PSF on reaction times (for detailed results see the table below).

[21] English always presents inflectional morpheme in word-final position, similarly to Czech, which is, however, morphologically richer and has a substantially broader variety of inflectional morphemes at its disposal (Šmilauer, 1973), along with inflectional endings for gender and person such as *-a* or *-o* (Melichar & Styblik, 2009: 118 - 120). The English inflectional system is thus consonantal in nature, while Czech combines both consonantal and vocalic changes. Czech L2 speakers can, nevertheless, be more sensitive to morphology in general since paying attention to inflectional endings in their L1 is an essential strategy to attain the correct meaning of the word – see, for instance, that the conditions are being distinguished at the A1 level at a pattern similar to the C1 level and even to the native speakers tested by Cilibrasi (2016).

	Df	Sum Sq	Mean Sq	F value	P-value
Level	5	0.097	0.019	0.791	0.555
PSF	1	0.320	0.320	13.035	**< 0.001** ***
Level:PSF	5	0.032	0.006	0.263	0.933

Table 7. The analysis of variance (ANOVA) table for the atomist analysis. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Interestingly enough, these results suggest that the performance is also influenced by the positional segment frequency. This shows that even in the sublexicon, the phonotactic probabilities of the items still play some role, indicating that lexical frequencies might have served as a predictor of the performance, that the participants were sensitive to phonotactic information and used it in the processing of spoken novel words. This suggests that morpheme decomposition might be simultaneously accompanied by some frequency effects and that both theories might equally predict the reaction times. The analysis of variance has also shown no significant effect of the level on this interaction, suggesting that all participants use this strategy of frequency effects similarly at all language levels. This is even more interesting given that the dataset includes data from the A0 level (that has zero or very little knowledge of English and of its frequency effects); however, the sample of the A0 level is too small to draw any specific conclusion. This finding merely opens the discussion of possible transfer effects from L1 Czech and offers ideas for future research in the area.

To see the strength of correlation between the reaction times and the positional segment frequency, mean reaction times for each stimuli item were calculated, and the values for each novel word were correlated with PSF taken from Cilibrasi (2016) (for the full list of novel verbs with the PSF values, see Appendix 1), using Pearson's correlation. The correlation between reaction times and PSF turned out significant, $r(118) = -0.316$, $p < .001$. The strength of the correlation is not very high, yet the results of Pearson's correlation were still significant. The figure below then provides graphical visualization of the effect of PSF on reaction times. The trendline in the plot nicely shows that, as predicted at the beginning of this section, the participants reacted quicker to items with higher phonotactic probabilities.

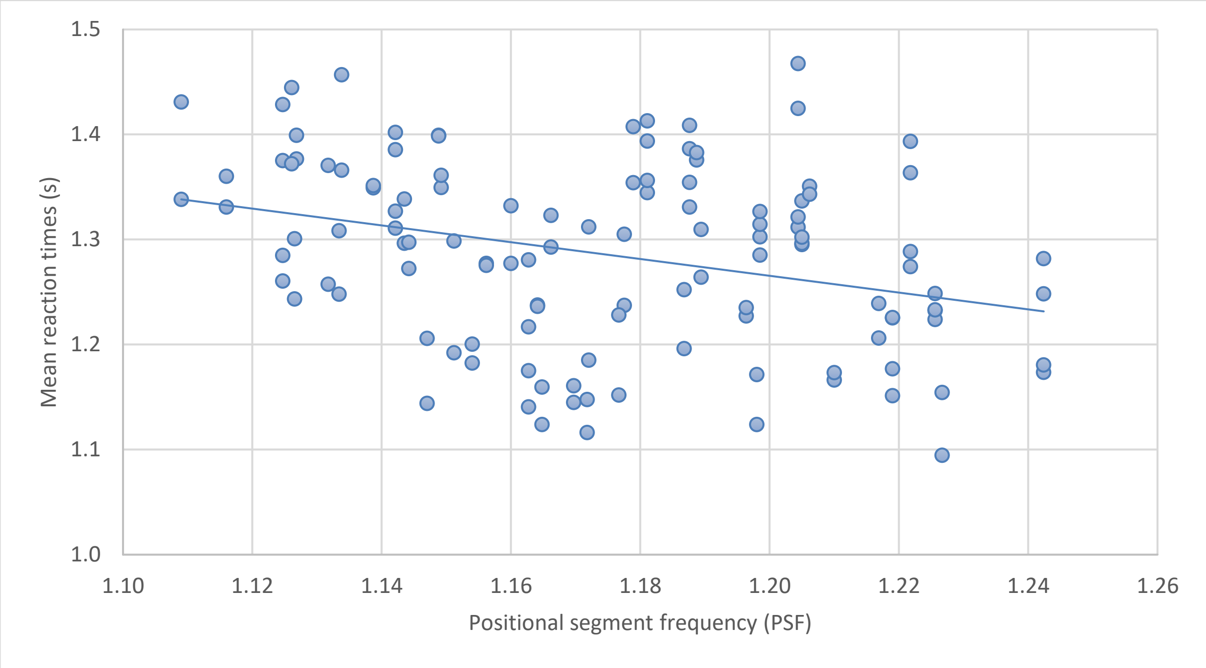

Figure 8. Mean item-based reaction times as a function of the stimuli item's positional segment frequency.

3.4.1.4 Summary of the reaction-times analysis

The morpheme-stripping analysis has shown that the A1 and A2 levels performed similarly to one another but differently from the B1, B2, and C1 levels, which, in turn, performed similarly to each other. We have also observed a gradual decline in the mean reaction times with higher proficiency (with the exception of the A0 level and the B2 level). The analysis has also shown that the participants reacted to different conditions differently, with the potentially morphological condition taking the longest to be distinguished, followed closely by the control and the non-morphological condition. All groups showed substantially the same pattern of reaction. The significant differences between the conditions indicate that morpheme stripping might be at play during the perceptual processing of the stimuli and that the participants might be, at the same time, analysing the phonological quality of the stem simultaneously with the affix.

The analysis also indicated that the stimuli type affected reaction times in such a way that participants reacted quicker when presented with the minimal pair consisting of identical items, while the reaction to the pair with different items took longer. The reaction times were also influenced by the duration of the presented stimuli pair, with a shorter novel word taking quicker to be discriminated. The significance of the interaction between the condition and the level suggests that participants at different language levels react differently to different conditions, with the A0 level not making a difference among the three conditions, while the rest of the language levels did. We have also noticed gradual speeding of reaction to the presence of morphology in the course of L2 acquisition. The significance of the interaction between the condition and the length has shown that the morphological condition is the condition that is the most sensitive towards stimuli duration (based on the apparent steepness of the trendline), followed by the control, and lastly by the non-morphological condition; yet with longer-duration novel words, the morphological conditions still took the longest to be discriminated.

The unit-shoring analysis has shown a significant effect of the positional segment frequency on reaction times in all participants, suggesting frequency effects are used similarly at all language levels. The graphical visualization has also shown that the participants reacted quicker to items with higher phonotactic probabilities, which corresponds to the beliefs of the atomist theory of inflection storage. The two analyses have thus given us evidence for both the decompositional hypothesis (based on the apparent difference in reaction times between the conditions; we have to mind the confound with stimulus duration, though) and the atomist hypothesis (based on the significant effect of the positional segment frequency on the quickness of reaction).

3.4.2 Accuracy analysis

The accuracy analysis attempts to answer the question of whether the correct recognition of a given stimulus is influenced by the type of the minimal pairs, their condition, or duration, and whether this performance differs at different language levels. The findings would offer us further information on the processing of English inflectional morphology.

A total of 10,986 answers were recorded (448 for the A0 level; 1,825 for the A1 level; 1,709 for the A2 level; 2,302 for the B1 level; 2,366 for the B2 level; and 2,336 for the C1 level). The mean proportion of correct answers for each condition (morphological, non-morphological, control) and type (minimal pairs with identical or different items) divided by the language level is shown in the table below (notice that all language levels performed at ceiling).

	N	Condition 1 (morpho)	Condition 2 (non-morpho)	Condition 3 (control)	Type "same"	Type "different"
A0 level	448	95%	91.5%	80%	97%	80%
A1 level	1,825	99%	96.5%	95%	98%	96%
A2 level	1,709	98%	98%	95%	97%	97%
B1 level	2,302	97.5%	98%	95%	97%	96.5%
B2 level	2,366	97%	98%	97%	97.5%	97%
C1 level	2,336	98%	98%	95%	98%	97%
All part.	10,986	98%	97.5%	95%	97.5%	96%

Table 8. Descriptive statistics of the dataset used for the accuracy analysis.

For each trial, key presses were recorded (along with the ID number, the language level, the placement test score, the type (same vs. different), the condition (morphological, non-morphological, control), and the length of the minimal pair (in seconds).

To investigate whether the data showed significant effects of various independent variables on the chosen answer, the results were analysed with a generalized mixed-effects model.[22] The dependent variable used was the accuracy (i.e., whether the answer was correct or not) to presented stimuli; the independent variables (fixed factors) included (a) the level, (b) the condition, (c) the type, and (d) the length (duration) of the stimuli. Once again, the independent variables were imported into the model simultaneously. Participants and the actual novel words were then used as random effects (the simplest and most-frequently-used form of random effects as described above in the reaction-times analysis).

[22] The data meet all the necessary assumptions of the model.

3.4.2.1 Initial data analysis

Again, we followed Pérez et al. (2016)'s procedure to find the most precise model for the combination of our independent variables, using the random effects for participant and item (see (1|part) + (1|item) above). Our optimization started with the highest number of possibly interacting independent variables: the level, the condition, the type, and the length, and it was compared with simpler interactions (see Pérez et al., 2016). Since the deletion of the four- and three-way interactions did not give a statistically significant outcome ($p = 0.69$ and $p = 0.34$ respectively); the four- and three-way interactions were removed from the model, followed by the interaction between the condition and level ($p = 0.132$), condition and length ($p = 0.536$), type and length ($p = 0.513$), and between the level and length ($p = 0.893$). We kept only the interactions between the condition and type ($p < .001$ ***) and the type and level ($p < .001$ ***) whose removal turned out significant. Since the removal of the length from the model did not result in a significant result of the ANOVA ($p = 0.997$), the final model included the level, the condition, and type as the combination of independent variables and participants and items (novel words) as random effects, resulting in this R code: cond:type + type:level + cond + type + level + (1|part) + (1|item). To account for possible multicollinearity of the independent variables, the variance inflation factor (VIF) was checked for each model. The values for the analysis are as follows: condition - 1; type - 1; level - 1.

First, an analysis was run on the whole dataset to see if participants at different levels performed differently and whether there was any effect of the condition, type, or the level (and, therefore, if individual post-hoc tests were needed). The overview of the linear model can be seen in Appendix 1. The subsequent analysis of variance showed a significant p-value for the condition, type, level, the interaction between the condition and type, and between the type and level (for detailed results, see the table below):

	Df	Sum Sq	Mean Sq	F value	P-value
Condition	2	34.20	17.10	**17.10**	**< .001*****
Type	1	10.27	10.26	**10.26**	**.002 ****
Level	5	17.31	3.46	**3.46**	**.007 ****
Condition:type	2	19.89	9.94	**9.94**	**< .001*****
Type:level	5	20.29	4.06	**4.06**	**.002 ****

Table 9. The analysis of variance (ANOVA) table for the accuracy analysis. The numbers in bold indicate a significant effect. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

The results show a significant effect of the language level, indicating that participants at different levels perform differently to the presented stimuli, with different levels of accuracy. The findings also show that participants reacted differently to the three conditions, which hints at possible effects of morphology on the answer. The results also suggest that the answer was influenced by the stimuli type, suggesting that the participants were sensitive to the type of the minimal pair when pressing the key. The significance of the interaction between the condition and the type suggests that participants reacted differently to each stimuli type in the three conditions. The significant effect of the interaction between the type and level also suggests that the accuracy of the answer to the stimuli of different types varies among the language levels. This provides us with enough evidence to perform group post-hoc tests and see how the results differ for each stimuli type, while the rest of the findings asks for graphical visualization to make us understand the relationship among the variables.

3.4.2.2 Group post-hoc tests

3.4.2.2.1 The role of the level

The initial ANOVA has shown a significant effect of the language level on the accuracy, which suggests that participants at different levels provide answers with different levels of accuracy to different stimuli. The figure below shows the graphic visualization of the participants' performance at different language levels. The plot shows an evident difference in accuracy between the A0 level and the rest of the proficiencies, while the A1 to C1 levels seem relatively homogenous at first glance (with a slight progression to higher accuracy with higher proficiency).

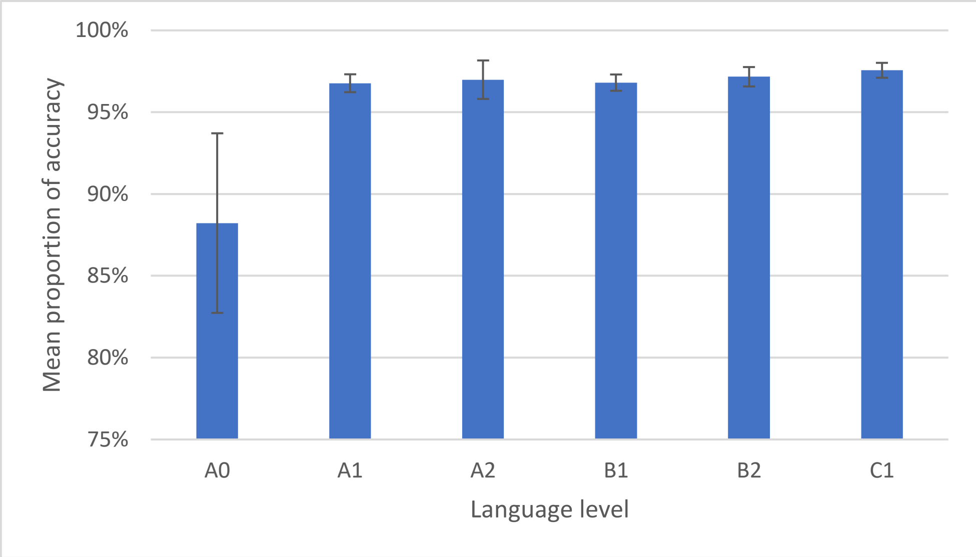

Figure 9. Mean proportion of correct answers at each language level.

To see how accuracy varies among the language levels, independent samples t-tests were run among individual language levels. The summary below shows that, as predicted from the plot above, the accuracy of the A0 level is significantly different from all other language levels. All the other interactions turned out not significant, indicating that the rest of the language levels performs with similar accuracy.

	A1	A2	B1	B2	C1
A0	t (119) = -4.59, **p < .001**, two-tailed	t (119) = -4.75, **p < .001**, two-tailed	t (119) = -4.65, **p < .001**, two-tailed	t (119) = -4.94, **p < .001**, two-tailed	t (119) = -5.09, **p < .001**, two-tailed
A1		t (119) = -0.46, p = .65, two-tailed	t (119) = -0.11, p = .91, two-tailed	t (119) = -0.94, p = .35, two-tailed	t (119) = -1.39, p = .17, two-tailed
A2			t (119) = 0.36, p = .72, two-tailed	t (119) = -0.43, p = .66, two-tailed	t (119) = -0.90, p = .37, two-tailed
B1				t (119) = -0.85, p = .4, two-tailed	t (119) = -1.32, p = .19, two-tailed
B2					t (119) = -0.52, p = .60, two-tailed

Table 10. Results of the post-hoc t-tests among individual language levels. Significant results are indicated in bold.

The low performance of the A0-level group is expectable since the non-familiarity with the target language seems to be playing a role here. However, the size of the group (only 4 participants) is too small to draw any specific conclusions about the performance, let alone about possible transfer effects.

3.4.2.2.2 The role of the condition

The initial analysis has also shown a significant effect of the condition on the accuracy, suggesting that participants provide answers of different accuracy for different conditions. The interaction among the conditions was first plotted in the figure below.

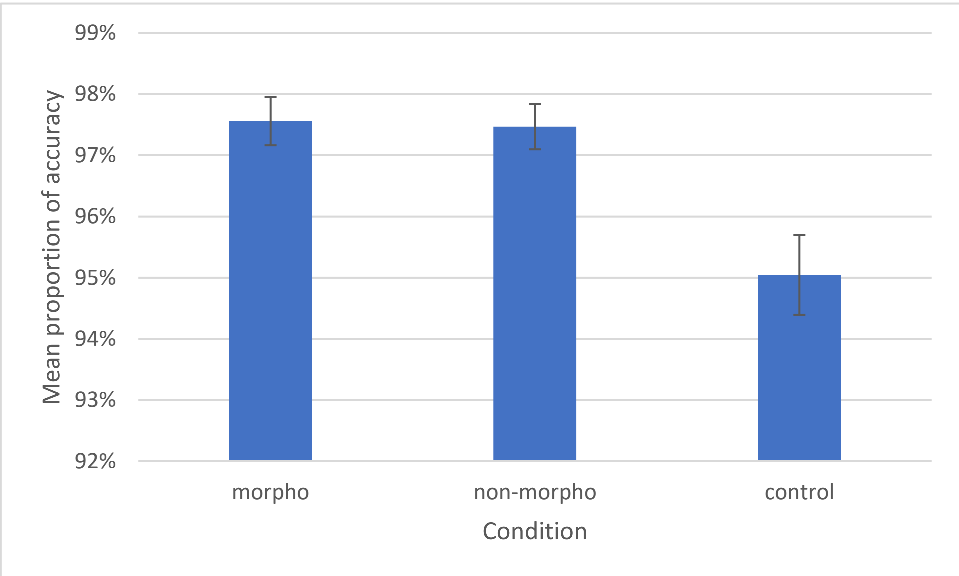

Figure 10. Mean proportion of correct answers in the three conditions (morphological, non-morphological, and control conditions).

The plot shows that out of the three conditions, the morphological condition produced the highest accuracy, followed closely by the non-morphological condition, and lastly by the control condition – this applies to all language levels since the initial ANOVA has shown no interaction with the language level. This visualization suggests that the morphological condition with potential morphosyntactic information is the easiest to distinguish correctly, while the control condition (which follows the rules of English morphophonology but could never be morphological) is the least accurate.

Post-hoc t-tests were then run among individual conditions to see if their difference turns out significant or not. The results (summarized in the table below) show that the accuracy in conditions 1 and 2 is similar (no significant difference is detected between them), but conditions 1 and 2 differ in accuracy from condition 3.

	cond 2 (non-morpho)	cond 3 (control)
cond 1 (morpho)	t (39) = 0.16, p = .87, two-tailed	t (39) = 3.29, **p = .001**, two-tailed
cond 2 (non-morpho)		t (39) = 3.22, **p = .002**, two-tailed

Table 11. Results of the post-hoc t-tests among individual conditions. Significant results are indicated in bold.

Since the control condition turned the least accurate, we can speculate about the participants' sensitivity towards the presence of morphological information, their awareness of English morphophonological rules, and its effect on accuracy. The control condition seems to be the most cognitively demanding as it produced the lowest value of accuracy out of the three conditions, suggesting that its phonological endings that can never be morphological in English (even in a different context as is the case in the non-morphological condition) are the hardest to process, probably due to the lack of familiarity. The control condition is also the only condition that ends in a phoneme that cannot be morphological in any given context. The fact that all language levels reacted with the least accuracy to this condition thus suggests that the bound morphology might have been surprising for the participants or somehow less clearly distinguishable from the other conditions. However, the true nature of this lowered accuracy can only be hypothesized at the present moment.

3.4.2.2.3 The role of the type

The initial analysis has also shown a significant difference in the participants' accuracy when presented with minimal pairs of identical and different items. The pairwise comparison of accuracy to the two stimulus types has shown a significant difference between the accuracy in each stimuli type (t (59) = -2.32, p = .02, two-tailed), with the "same" type showing higher accuracy. The figure below depicts a graphical visualization of the performance.

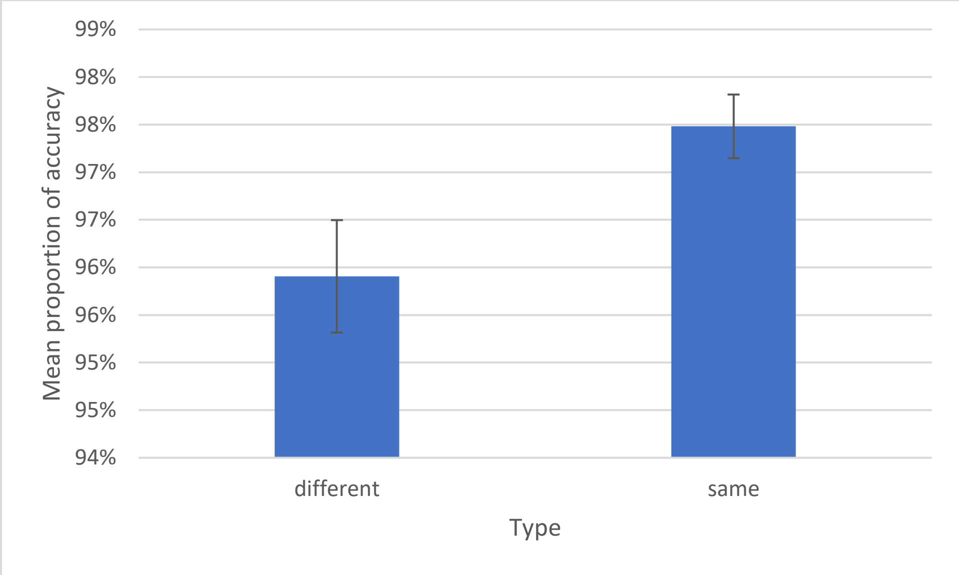

Figure 11. Mean proportion of correct answers based on the type of the presented stimuli.

The plot clearly shows that the participants produced more correct answers when presented with the minimal pair consisting of identical items, while the accuracy of the minimal pair with different items is significantly lower. The higher accuracy with the "same" type could be explained by the general fact that recognising a small difference in two similar items is more cognitively demanding than recognising two items as being identical, analogically to the type effect found in the reaction-times analysis (see above). The contrast in the final word-position in the "different" type is then more cognitively demanding and produces less accurate answers.

3.4.2.2.4 The role of the interaction between the condition and type

The next significant interaction of independent variables in the initial analysis was the interaction between the condition and the type, which suggests that participants provide answers of different accuracy to novel verbs of different stimulus types in different conditions. Post-hoc t-tests were, therefore, needed to understand the performance in greater detail. For these purposes, a data subset was created for each condition, and independent samples t-tests were run between the two stimulus types. The results of the post-hoc t-tests can be seen in the table below:

	same vs. different type
cond 1 (morpho)	t (19) = -0.62, p = .53, two-tailed
cond 2 (non-morpho)	t (19) = 0.74, p = .46, two-tailed
cond 3 (control)	t (19) = -3.92, **p < .001**, two-tailed

Table 12. Results of the post-hoc t-tests among individual conditions with a focus on the potentially significant difference of the type. Significant results are indicated in bold.

The findings of the post-hoc t-tests indicate that the participants displayed no significant difference in accuracy between the two types in the morphological and non-morphological conditions. Their performance, therefore, was not influenced by the identity of the items in the presented stimuli. However, in the control condition 3, there was a significant difference between accuracy for type "same" and "different." The figure below shows how the participants' accuracy changed with the stimulus type in the control condition.

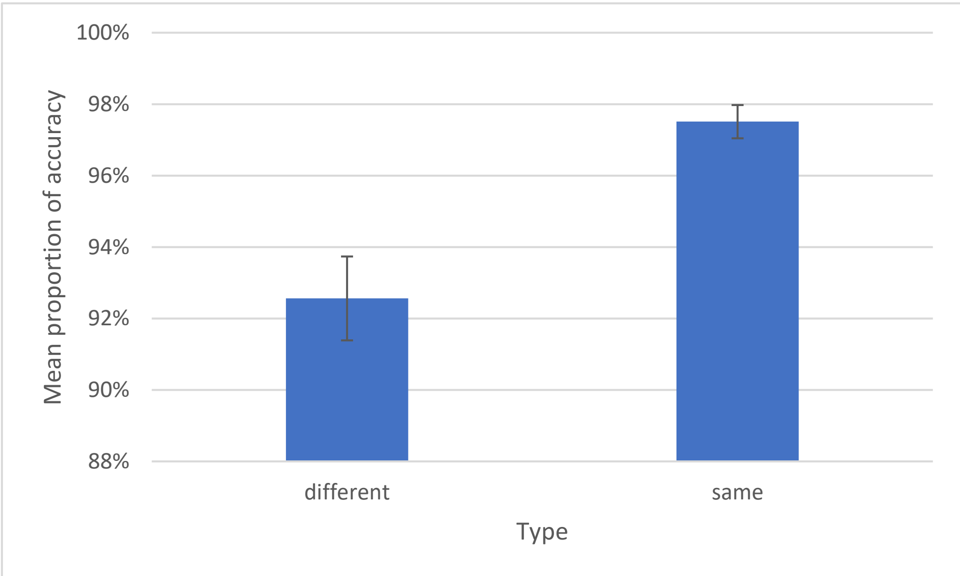

Figure 12. Mean proportion of correct answers based on the type of the presented stimuli in the control condition 3.

As with the role of the type discussed above, the plot clearly shows that the participants produced more correct answers when presented with the minimal pair consisting of identical items, while the accuracy with the pair with different items is significantly lower. The higher accuracy with the "same" type could be again explained by the lower cognitive demand of the task. The lowered accuracy in the "different" type could then be linked back to our hypothesis about the unfamiliarity of the control condition and its allomorphic endings discussed above.

3.4.2.2.5 The role of the interaction between the type and level

The last significant interaction of independent variables found in the initial analysis was the interaction between the type and the language level, which suggests that participants at different language levels provide answers of different accuracy to novel verbs of different stimulus types. Again, post-hoc t-tests were needed to understand the difference in performance. A data subset was created for each language level, and independent samples t-tests were run between the two stimulus types. The results of the post-hoc t-tests can be seen in the table below:

	same vs. different type
A0	t (59) = 5.22, **p < .001**, two-tailed
A1	t (59) = -0.85, p = .4, two-tailed
A2	t (59) = -0.26, p = .8, two-tailed
B1	t (59) = -0.64, p = .53, two-tailed
B2	t (59) = -0.53, p = .6, two-tailed
C1	t (59) = -1.95, *p = .06*, two-tailed

Table 13. Results of the post-hoc t-tests among individual language levels with a focus on the potentially significant difference of the type. Significant results are indicated in bold, marginal in italics.

The findings of the post-hoc t-tests have shown that only at the A0 level do participants produce differently accurate answers based on the stimuli type. At the C1 level, the difference between the two types was only marginally significant when two-tailed. The rest of the language levels made no difference between the identity of the stimuli when choosing the answer to the presented novel verb. The performance of the language levels was plotted to see how accuracy changes with the stimulus type (see the figure below).

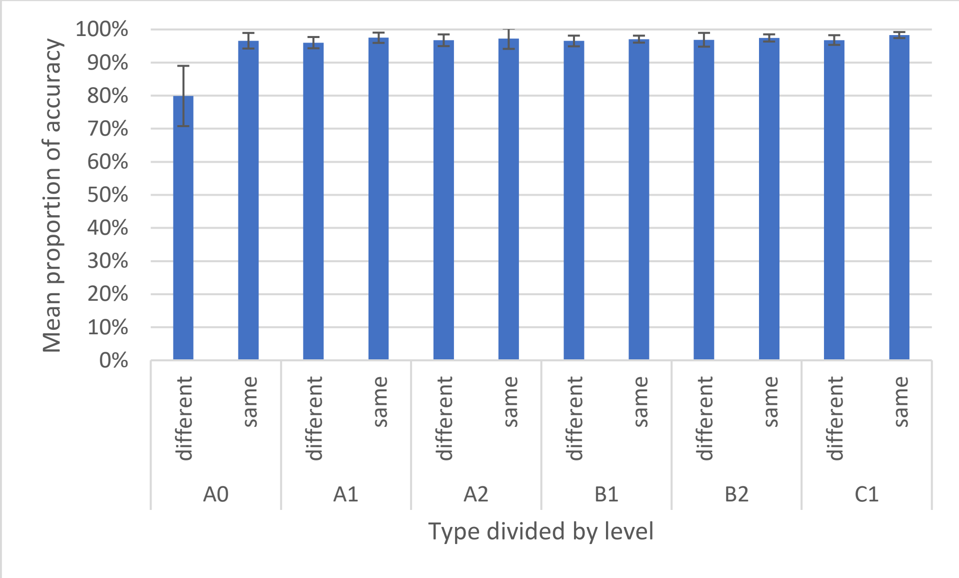

Figure 13. Mean proportion of correct answers based on the type of the presented stimuli at individual language levels.

Similarly to the previous visualization of accuracy based on type, the plot clearly shows that the participants at each level produced more correct answers when presented with the minimal pair consisting of identical items, while the accuracy with the pair of different items is lower.

3.4.2.3 Summary of the accuracy analysis

The initial analysis of variance showed that in terms of accuracy, the A0 level performed in a significantly different way from all other language levels, producing less correct answers to the presented stimuli. All other language levels performed with similar accuracy. The findings also show that participants reacted to the three conditions differently, with the morphological condition producing the highest accuracy, followed very closely by the non-morphological condition and only then by the control condition. The same pattern was observed in all groups. The morphological condition seems to be the easiest to distinguish correctly, while the control condition seems to be the most cognitively demanding.

The results also suggest that the answer is influenced by the stimuli type, with the participants producing more correct answers when presented with the minimal pair consisting of identical items. The higher accuracy with the “same” type could be explained by the lower cognitive demand of the task. Further analysis run on the significance of the interaction between the condition and the type has shown that only in the control condition do participants answer differently to the novel verbs of the two stimulus types, with higher accuracy for the "same" type. The significant effect of the interaction between the type and level has also shown that only at the A0 level do participants produce differently accurate answers based on the stimuli type, again, with the "same" type responses producing higher accuracy. The rest of the language levels makes no difference between the identity of the stimuli when choosing the answer to the presented novel verb.

3.5 Discussion

3.5.1 Summary of the findings

Our experiment extends previous findings on monolingual speakers to second language learners. The morpheme-stripping sub-analysis in the reaction-times analysis has shown a significant difference between the A1 and A2 levels and the B1, B2, and C1 levels, with a gradual decline in mean reaction times with higher proficiency (with the exception of the A0 level and the B2 level). The analysis has also shown a significant effect of condition on reaction times, with the potentially morphological condition taking the longest to be discriminated, followed closely by the control and the non-morphological condition. This result is in line with previous studies on monolingual speakers, such as Post et al. (2008) and Cilibrasi et al. (2019). The participants also reacted quicker to the minimal pair consisting of identical items, while the reaction to the pair with different items took longer. The length of the stimuli also played a role in reaction times, with shorter novel words taking quicker to be discriminated. Further analysis has also shown that the morphological condition is the condition that is the most sensitive towards stimuli duration and that there is gradual speeding of reaction times in the presence of morphology in the course of second-language acquisition. This finding partly reduces the strength of any conclusion drawn from the condition effect since this very fact suggests that there is a confound between morphological complexity and duration effects, even though the presence of independent condition and duration main effects suggest a genuine presence of morphological processing in inflected novel words. The results have also shown that the A0 level failed to discriminate differences between individual conditions, while the rest of the proficiency groups retained the pattern described above. Some (at least basic) knowledge of English might, therefore, be needed for properly distinguishing the presence of bound morphosyntactic information. In addition, the unit-storing sub-analysis has shown a significant effect of the positional segment frequency on reaction times in all participants, suggesting frequency effects are used similarly at all language levels, once again weakening a purely morphological explanation. The participants reacted quicker to items with higher phonotactic probabilities, which corresponds to the core claims of the atomist theory of inflectional processing.

The accuracy analysis has shown that the A0 level produced significantly less correct answers to the presented stimuli than any other language level. The findings also show that the morphological condition produced the highest accuracy, followed very closely by the non-morphological and the control condition. This pattern may be due to the fact that the novel verbs in the control condition ended in phonemes that cannot be morphological in any given context, sounding unfamiliar to the listener. The answers have also shown the influence of the stimuli type, with the participants producing more correct answers when presented with the minimal pair consisting of identical items. Only in the control condition did participants answer differently to the novel verbs of the two stimulus types, with higher accuracy for the "same" type. Also, at the A0 level (and marginally also at the C1 level), the participants produced differently accurate answers based on the stimulus type, with the "same" type responses producing higher accuracy.

3.5.2 Discussion and conclusion

These findings contribute to the body of research on morphological processing in novel words. Recent perception studies at the sublexical level using novel verbs (e.g., Post et al., 2008; Cilibrasi, 2016; Jiránková, 2017; Cilibrasi et al., 2019) have provided some evidence that inflectional morphemes can be detected sublexically and that both monolingual English speakers and more proficient L2 speakers of English are sensitive to the morphosyntactic information in novel words even at the sublexical level in the absence of any lexical meaning of the word. These previous studies have also shown that the participants seem to analyse the phonological quality of the bound morpheme together with the quality of the stem. The current study shows that these findings might also be present but less clear cut in L2 speakers since their performance seems to be influenced not only by the presence of morphology but also by the stimuli length and phonotactic probabilities of the presented stimuli.

All considered, our findings suggest that the perception of English inflectional morphology might be explained by the morpheme-stripping (or morpheme sensitivity) hypothesis, but it also suggests that some frequency effects are parallelly active in L2 learners (see Jiránková, 2017). So far, evidence for the possible effect of morpheme-stripping in the sublexical perception of inflectional morphology has been observed mostly in English native speakers (Post et al., 2008; Cilibrasi, 2016; Cilibrasi et al., 2019), but also in higher-proficiency second-language learners of English (Jiránková, 2017). Our study used a perceptual discrimination task to investigate whether evidence for morpheme stripping (or sensitivity) or unit storing extends to all L2 learners of English at various stages of L2 acquisition and Czech native speakers with no or very little knowledge of English and whether the process behind the perception differs among individual levels, starting with a proficiency sample at the very beginning of morphological acquisition and moving to the language sample closest to English native speakers.

Our data suggest that all analysed language levels except for the A0 level might be decomposing inflectional morphology into stems and affixes during perception since the participants took significantly longer to discriminate novel words with potential morphosyntactic information than those without it. The significant difference in reaction times to novel words with potential morphosyntactic information and those without cannot be explained only phonologically, suggesting that morphological decomposition might have been in process. However, the length effect partly undermines this claim because the participants were the most sensitive towards the length of the presented stimuli in the morphological condition, suggesting that not only the morphosyntactic information but also the stimuli length affects the performance. Further analysis has also shown a significant effect of the positional segment frequency on reaction times, suggesting that the quickness of reaction was not dependent on the morphosyntactic information or length only, but frequency effects were playing a significant role as well.

The results have also confirmed that morphological decomposition might not be blind and that L2 learners might also implicitly analyse the phonological quality of the stem and the affix simultaneously, suggesting that morphology and phonology might go hand in hand in morpheme stripping and word parsing. These findings are consistent with previous research done, for instance, by Pater (2006), Harris (2011), or Cilibrasi (2016). Particularly, Grainger and Ziegler (2011) proposed the existence of "a fine-grained parser" who is capable of detecting specific grapheme groups that may carry morphological information and do so without accessing the meaning of the word itself.

Our findings suggest that a similar strategy is used by L2 learners during aural language perception (comparable to what Cilibrasi (2016) suggested for L1). Our participants might use morphological decomposition and stem analysis from the very beginning of L2 acquisition, implicitly and apparently as a result of automatic language processing. They might do so in the absence of any lexical word meaning, which suggests that even for L2 learners of English morphological decomposition and stem analysis might occur sublexically, similarly to English native speakers studied, for instance, in Post et al. (2008), Cilibrasi (2016), or Cilibrasi et al. (2019).

A significant difference has also been found between the control and the non-morphological condition. This finding cannot be explained purely by the presence of morphology similarly to the data above since none of those conditions is morphological per se. The reason for the contrast between the two conditions is, therefore, likely phonological in nature since the non-morphological condition presents a voicing contrast, while the control condition does not. If we followed this line of argumentation, we could explain the difference between the morphological and the non-morphological condition also by the difference in voicing (and the control condition was involved in the first place to avoid this kind of explanation), which would undermine the previous explanation of the data. However, the fact that the morphological condition takes longer than the phonological control condition suggests that voicing cannot be the only reason but only a concurrent one. As of now, we are not fully capable of explaining the difference based on any previous research done in the area. Neither are we capable of guessing what effect exactly both morphology and phonology have on this difference between the morphological and non-morphological conditions, especially when lengthening (one of the properties of voiced consonants) appears to have a significant impact on reaction times per se. In other words, while we do certainly observe that morphological items are slower to be processed, we are not able to fully explain why.

Even though the data showed a significant difference between the A1 and A2 levels and the B1, B2, and C1 levels with a progression in the quickness of reaction with higher proficiency (except for the A0 and B2 data), the pattern across conditions was the same for each language group: morphological > control > non-morphological. Interestingly enough, the same differences were also found in English native speakers tested in Cilibrasi (2016). English proficiency led to quicker reaction times, but it did not generate any significant difference in the reaction to the three conditions. L2 learners' processing of inflectional morphology is therefore comparable to that of native speakers. The lack of proficiency effect was also found in the use of morphological decomposition and stem analysis since all language levels (including the lowest ones) performed in the same way. Looking at a possible explanation of these findings, we can refer back to the one suggested in Jiránková (2017). First, the explicitness of grammar instruction at Czech schools might play a role since inflectional morphemes are introduced very early on and rigorously drilled. Second, the morphological richness of Czech might be a prominent contributor to such behaviour in the perception of English inflectional morphology since paying close attention to Czech inflections is crucial for understanding the meaning of a sentence (given the lack of fixed word order in Czech), and this approach may be replicated by Czech speakers when they process English-like items. This explanation is supported by previous research. Studying the difference between English and Chinese, Ku and Anderson (2003) showed that L2 learners with a morphologically rich L1 are in general more sensitive to the morphological processes that are active in their native languages. Their Chinese participants lagged behind the English subjects in the use of derivational morphology, and the authors hypothesized that the reason might be that Chinese word-formation is not very rich. Czech speakers thus may pay closer attention to word endings and transfer this strategy to English, using it implicitly since the participants were not told beforehand what the goal of the testing was, and the speed of the task also deems any conscious rule application or internal reflection rather unlikely. The effects observed in the task may have been, therefore, a result of unconscious automatic processing, as proposed already in Jiránková (2017). On the other hand, while the overall pattern at the A1 level is comparable to what is found in the other groups, proficiency does play a role in that beginners are overall significantly slower and significantly less accurate in the tasks.

Our data also confirmed that phonotactic probabilities play a decisive role in the quickness of reaction. Positional segment frequency had a significant effect on reaction times irrespective of the levels, and it also positively correlated with reaction times to individual novel words. This clearly shows that even at the sublexical level, some frequency effects and analogical processing of novel words are still active and may be just as important in determining the reaction times as the morphological decomposition.

Our findings, together, suggest the complex interaction of morpheme stripping and phonotactic probabilities effects.[23] Importantly, our data suggest that morpheme decomposition may be applicable in the absence of word meaning as well. This claim extends previous support of morpheme-stripping processes associated with morphological perception (e.g., Post et al., 2008; Cilibrasi, 2016; Jiránková, 2017; Cilibrasi et al., 2019) to (less proficient levels of) second-language learners of English. All the previously-mentioned studies use their findings to demonstrate that the perception of English inflectional morphology might be best explained in terms of the dual-route model, showing decomposition effects (e.g., Post et al., 2008; Cilibrasi, 2016; Cilibrasi et al., 2019) parallel to frequency effects in L2 speakers (e.g., Jiránková, 2017). The current study shows that these claims could also be extended to lower-proficiency L2 learners, though with important limitations related to duration effects. Our findings also bright further insights into how morphologically-inflected forms are stored, suggesting that both morphological decomposition and analogical processing (based on the significance of phonotactic probabilities) might be active in the sublexicon of L2 speakers. This finding seems to suggest that both hypotheses might be, in fact, valid and applied in aural language processing. One possibility is that the system might be simply redundant: the hearer might rely both on the morphological structure of the word but also on its phonotactic probabilities and on how the word functions as a phonological unit to process it. Both the use of rules and analogy might be, thus, applied simultaneously. The two suggested hypotheses of perceptual processing of inflectional morphology usually make a thick dividing line between decomposition and unit

[23] The division into decompositional and atomist approaches is only one of the theories of morphological processing. We may also like to mention the entropy theory (e.g., Lõo et al., 2018) – an approach that accentuates the need to integrate the studied morphological form within its morphological paradigm that fulfils a specific morphological function (e.g., past-tense morpheme possibilities). This morphological paradigm and its frequencies in L1 then may determine the learner's sensitivity to certain L2 morphological features. In their paper on Estonian inflections, Lõo et al. (2018) investigated how "inflectional paradigm size and morphological family size affect production latencies and articulation durations" (Lõo et al., 2018: 71) of participants producing Estonian nouns inflected for case, and they found that the size of the inflectional paradigm does predict word naming latencies and acoustic duration. The entropy values for each morpheme in this functional paradigm in L1, i.e., the frequency or probability of its distribution, can influence its perception and production in L2, and the general sensitivity to morphology transferred from a morphologically rich L1 (as suggested, for instance, by Ku and Anderson) may not be the only predictor in L2 performance.

storage, trying to distinguish the former from the latter. However, given the results of our research that suggest that morpheme stripping might be accompanied by some frequency effects, and in line with previous research based on two seemingly incompatible models (e.g., Cilibrasi et al., 2019), this discussion attempts to propose a less diving conclusion. This redundancy of two mechanisms operating in parallel has already been proposed by redundancy models. Schreuder et al. (1999), for instance, suggest that the two mechanisms can be activated at the same time, each for its own domain (e.g., frequent and infrequent items – see Chapter 2). Using Schreuder et al.'s proposal, we can justify the possible simultaneous activation of morphological decomposition and frequency effects in our data.

In conclusion, this first experiment has allowed us to take a closer look into the mechanisms that underlie the perception of inflectional morphology in L2 learners of English, showing that morphological decomposition might be accompanied by frequency and length effects in the sublexicon and a more embracing explanation of such perception is needed.

4. EXPERIMENT 2: THE PRODUCTION EXPERIMENT[24]

4.1 Introduction

The baby likes to dize. Look, there it is dizing. Every day it dizes. So yesterday it…dized? Or *doze*? The answer to this question poses a number of difficulties. The discussion of whether novel morphologically-inflected forms are produced by the use of rules (e.g., Prasada & Pinker, 1993) or by analogy based on stored examples (e.g., Bybee & Slobin, 1982) is continuously addressed in the field, and attention is being paid especially to the difference between regular and irregular inflection, particularly that of the English past tense, due to its relatively clear-cut distinction between the two forms. As such, English past tense suits perfectly the purposes of adjudicating between those two approaches.

Regular past-tense inflection realized by the "add *-ed* to the verb stem" rule (as in *jump – jumped* or *cook – cooked*) applies to thousands of verbs and is often productively used in the generalization of new word-forms (as in *zib – zibbed* in Berko, 1958), while irregular past-tense forms (such as *fall – fell*) apply to circa 200 known irregular verbs and often fall victim to overgeneralization by children (as in **sitted* instead of *sit*) (Pinker & Ullman, 2002). These irregular forms have been traditionally assumed to be acquired and stored as compact units in our memory (Prasada & Pinker, 1993). However, the simple idea that regular forms are obtained from rules and irregular forms consist of rote storage has proved problematic in several aspects – irregular verbs, for instance, form families of similar verbs such as *cut* and *put* whose past tense remains the same as the present form and thus allows for analogical generalization of novel verbs (see, for instance, Bybee & Moder, 1983). English also includes many quasi-regular inflections (Cilibrasi et al., 2019). Such past-tense forms are "obtained following a productive pattern, that is, however, less frequent than that normally labelled as regular" (Cilibrasi et al., 2019: 750). Examples of quasi-regular past-tense forms would be *feel/felt, keep/kept*, or *build/built*. The main question of the above-mentioned debate thus centres around whether "regular and irregular past tense forms are generated by two qualitatively distinct mechanisms or whether all forms are produced in a single, associative process" (Westermann & Ruh, 2012: 649). Despite the length of the debate, the question remains unanswered, and the views differ

[24] Parts of this chapter will be published (with some changes) in *Linguistics Pragensia*. The paper was in press at the time of the thesis submission.

diametrically. Generativist theories often attribute generalization to rules, connectionist theories attribute it to the association based on analogy, while hybrid theories attribute regular generalization to rules and irregular generalization to analogy, but they also suggest that frequent regulars verbs may be stored in memory with their inflection (Prasada & Pinker, 1993).

As a consequence, two models have been proposed for past-tense production: (i) the single-route model (see, for instance, Bybee & Moder, 1983), which posits that all past-tense forms are produced analogically to word forms that already exist in our mental lexicon, and (ii) the dual-route model (e.g., Prasada & Pinker, 1993), which assumes that past-tense forms are generated by two different processes: Regular past-tense forms are produced by a basic rule for past-tense inflection (i.e., adding *–ed* to the verb stem), while irregulars are stored as units and used in analogical generalization (Blything et al., 2018). It is evident that both models assume that the same mechanism is used in the generation of irregular past-tense forms (Blything et al., 2018). Investigating irregular verbs, therefore, does not help us decide which model is a better description of how past-tense forms are produced, and regular inflection needs to be studied instead.

The majority of psycholinguistic studies generally uses novel words to test morphological productivity (e.g., Prasada & Pinker, 1993; Albright & Hayes, 2003; Ambridge, 2010; Blything et al., 2018) since such a procedure requires generalization and avoids effects related to the lexical access (Blything et al., 2018). This assumption is also at the foundation of studies such as Albright and Hayes (2003) and Blything et al. (2018), who use an elicitation production task to examine which of the two models is better at describing morphological productivity in native speakers of English. This experiment builds on these two previous studies and aims to additionally explore (i) the development of morphological productivity of second-language learners of English and (ii) the effect of their L1 (Czech) on the L2 production.

4.2 Theoretical background and hypotheses

4.2.1 Theoretical background

4.2.1.1 Previous novel-verb studies

One of the pioneers in the field of novel-verb studies on morphological production is indisputably Jean Berko Gleason and her wug test from 1958. In this study, novel words were used to test how children acquire morphological rules. Children were shown pictures of imaginary creatures or activities and asked to supply past-tense forms of novel verbs (e.g., *zib*) or plural forms of novel nouns (e.g., *wug*) by completing a short statement related to the novel word. The experiment was designed to test the children's acquisition of morphological rules, and Berko used novel words on purpose since only then the child does not access the lexicon during the language processing, and the experiment can, therefore, shed light on the acquisition of rules. As Berko (1958: 150) explains, "we know that if the subject can supply the correct plural ending, for instance, to a noun we have made up, he has internalized a working system of the plural allomorphs in English and is able to generalize to new cases and select the right form." She then proceeds to give the example of *witch* and *witches*, a pair that the child might have simply memorized, and contrast it with the performance in her task: a child who forms the plural of the novel word *gutch* saying *gutches* has already internalized the rules for creating plurals, without anyone else teaching them (Rosenbaum, 2011), and this form cannot be retrieved from memory. Berko showed that very young children infer rules from the language around them, without relying on memorized units of stems with the inflectional morphemes (Karmiloff & Karmiloff-Smith, 2001) and that they are able to generalize them on new language material. Her findings suggest that in language production, children derive inflected forms using rules (this is further supported by over-regularization cases described in naturalistic recordings – e.g., forms such as **comed* or **hitted*, see, for instance, Marcus et al., 1992). The natural interpretation of this pioneering work of Berko (1958) is that verbs are (generally) inflected by the application of a rule.

This idea has been, however, criticized in other studies. Focusing on the storage of English verbal irregulars, Bybee and Slobin (1982) tested pre-schoolers, 8-10-year-olds, and adults, finding evidence for the storage of irregular past-tense forms as whole units in our mental lexicon, but also for the possibility of generalizing those forms based on schemas that describe their phonological properties, thus showing that "irregulars" are productive. The authors also proposed that the same lexical storage may be true even for regular past-tense forms (but lacked supporting evidence at that time). A year later, Bybee and Moder (1983) suggested that English native speakers treat irregular past-tense patterns productively based on a prototypical schema of similar stored irregular verbs.

Similarly, with their pattern-associator model simulating the learning process of a young child picking up English, Rumelhart and McClelland (1986) challenged the rule-based account of how a child acquires the English past tense by speaking in favour of implicit rules based on analogy. The findings of their study indicated that, contrary to previous empirical claims, a universal procedure is applied to both regulars and irregulars: The person is given the input stem, and "the resulting pattern of activation is interpreted as a phonological representation of the past form of that verb" (Bybee and Slobin, 1982: 38) based on how phonologically similar the base form is to the pre-stored past-tense forms, resulting in "on-line generalizations from the stored exemplars" (Bybee and Slobin, 1982: 38). Building on Rumelhart and McClelland's findings, Westermann and Ruh (2012) argued that all morphologically-inflected past forms are produced on the basis of a single mechanism based on analogy and association. All these studies can thus be considered as evidence for the existence of the single-route model in past-tense production.

Contrary to the preceding studies and in line with the dual-route model, Pinker (1991) proposed that while regulars are handled by a rule, irregulars are generated from our mental lexicon. In his opinion, default rules apply freely to all verbs unless the rule is blocked by a competing irregular form (Pinker, 1991). To evaluate how English native speakers produce novel morphologically-inflected words, Prasada and Pinker (1993) developed 60 novel verbs that varied in how similar they were to existing regular and irregular verbs and asked the participants to produce their past forms. Their findings have shown that the participants tended to inflect novel verbs that resembled irregular verbs as irregulars as long as their similarity to existing irregulars was relatively high; otherwise, the production of irregular past-tense forms decreased. On the contrary, the participants also produced a high amount of regularly inflected forms, and this production seemed independent of how much the novel verb was similar to existing regulars. Prasada and Pinker (1993) thus understood those findings as evidence for a hybrid, dual-route model of production in which "irregular forms are stored in an associative memory and regulars are produced by default irrespective of similarity to existing forms" (Westermann & Ruh, 2012: 657). In their view, the difference between regulars and irregulars lies in the fact that "regular inflection is a default process and applies to all forms unless explicitly blocked by an irregular stem" (Prasada & Pinker, 1993: 46, see also Pinker & Ullman, 2002). The strongest version of such a model maintains that "predictable inflected forms are never stored in memory," while the weaker version of the model admits that "prior storage of regulars is possible, and thereby might offer mild analogical assistance to the generalization of the regular inflection to similar forms" (Prasada & Pinker, 1993: 7). Looking at frequency effects, Alegre and Gordon (1999) showed that even regular forms could be produced as whole units when their frequency exceeds roughly 6 in a million words, while lower-frequency forms need to be composed through the use of a default rule, similarly to the weaker version of Pinker and Prasada (1993)'s model.

Albright and Hayes (2003), Ambridge (2010), and Blything et al. (2018) addressed these issues using acceptability judgement and novel-verb production tasks, testing adults and children with a clear intention to distinguish between the two suggested models. Albright and Hayes' study (2003) created novel verbs that showed phonological similarity or dissimilarity to existing regular and irregular English verbs. Using an elicited production task, the authors found that novel words that resembled existing regular verbs were more likely produced and judged as regular past forms by adult English speakers, and novel words that resembled existing irregular verbs were similarly more likely produced and judged as irregular past forms, in line with the single-route model. Replicating the judgement section of Albright and Hayes's study (2003), Ambridge (2010) tested children aged 6–7 and 9–10. Again, the study found that participants favourably judged novel irregular past forms that phonologically resembled existing irregular verbs. The older group of the English-speaking children more often accepted regular past forms of those novel verbs that highly resembled existing regular verbs, again in line with the single-route model. Blything et al.'s study (2018) builds on the two previous studies. Using the novel verbs created by Albright and Hayes (2003), the authors investigated the mechanisms that underlie morphological productivity, focusing on verbal morphology of the English past tense and recruiting groups of children from 3 up to 10 years of age. Their results have shown that "the likelihood of a novel verb being produced in regular past-tense form is positively associated with its phonological similarity to existing regular verbs" (Blything et al., 2018: 3), in line with the single-route model.

All three studies found that both regular and irregular forms of novel verbs are produced analogically to existing word forms stored in our mental lexicon and ruled by their similarity to these stored forms (Ambridge, 2019). These findings in the field of novel-verb studies provide us with tentative evidence that the single-route model might be better at describing the morphological production of novel forms in comparison to the dual-route model and, consequently, that analogy prevails over rules in regular past-tense production. The authors investigated both adults and older children of 9-10 years of age, and also smaller children at the peak rate of over-generalization, not only paying attention to mature linguistic systems but also to the systems still undergoing development (Blything et al., 2018).

4.2.1.4 Previous L2 studies

Previous studies on L2 learners seem to suggest that the dual-route model is not capable of fully explaining and describing the production of L2 learners. Studying how word frequency affects the production of regular verbs, Beck (1997) and Ellis & Schmidt (1998) reported findings inconsistent with the dual-route model. In their studies, the authors showed that frequency effects do not apply only to irregular verbs but also to regular ones. These results match those presented in Murphy (2000), who showed that the assumptions of the dual-route model are difficult to be applied to L2 data and are better explained by an alternative, associative model. Even though her L2 participants (similarly to native speakers) did add past-tense suffixes to novel regular verbs more than to the irregular ones, both native and L2 speakers used the novel verb's phonological resemblance to existing (ir)regulars as a decisive factor in producing past-tense forms – a finding not predicted by the dual-route model. These results thus undermine the claims of the dual-route mechanism that supports a complete dissociation of the processes behind the two past-tense forms and of the learning systems underlying language representation. Murphy (2000: 112) proposes "a more parsimonious account invoking a single set of associative learning mechanisms" instead. In more recent years, Cuskley et al. (2015) examined the difference between native and L2 production of past tense using a novel-word task, showing that both groups of participants show sensitivity towards phonological similarities between the novel and existing verbs (i.e., it is more probable that the participants will produce (ir)regular past-tense forms with novel verbs that phonologically resemble existing (ir)regular verbs). They have also noted that instead of showing a straightforward preference for the default regular past-tense rule, the L2 learners were generally more prone to using sub-rules when producing irregular past-tense forms. Similarly to Murphy (2000), Agathopoulou (2009) found that adult Greek L2 learners of English and English native speakers produced past forms of novel verbs that are not fully captured by the dual-route model (due to an apparent similarity effect found in the elicited forms) and noted no difference between how L1 and L2 English (ir)regular verb morphology is represented in the mental lexicon.

These studies offer interesting insights into morphological processing in L2 learners. The present research aims at exploring how these processing strategies are achieved by L2 learners of English with Czech as L1[25] and thus at describing the development of their L2 morphological system. Since Czech is a language with a complex and rich inflectional system, it creates an interesting contrast with the notoriously poor inflectional system of English and, therefore, provides a suitable ground for investigating possible transfer effects of Czech on the acquisition of English inflectional morphology. To achieve our aims, the study uses a production task with novel verb stimuli and tests participants with various proficiencies, including participants in the initial stages of learning, participants with a rather mature target language, and the range in between. Once again, the inclusion of a progression of language levels enables us to properly map the development of inflectional morphology in L2 learners and describe its specifics.

4.2.2 Hypotheses

Based on previous research both on monolingual and second-language production of morphologically-inflected forms outlined above, our tentative hypothesis was that the dual-route model would not fully explain the production of L2 learners and that native speakers would perform similarly to the participants tested in Albright and Hayes (2003) and Blything et al. (2018).

Regarding language proficiency, the thesis also operated with the tentative hypothesis that the lower levels might be more inclined to use rules due to lesser experience with L2, while the higher language levels might be more inclined to use analogy (and, therefore, resemble more native speakers of English) due to having more significant experience with the language and the analogical rules underlying past-tense formation.

[25] In comparison to the English past-tense morphology, Czech creates past-tense forms through the combination of the past stem of a verb and the addition of the so-called past-tense "*l*-forms," i.e., suffixes *-l, -la,-lo,-li/y*. These *l*-forms further attach suffixes for gender and number, offering a larger morphological variability in comparison to English inflection. In addition, number and person may be expressed by an auxiliary verb, which appears with the main verb and thus contributes to the morphological complexity of the sentence. Finally, the past stem can undergo stem alteration (Čechová, 1996), depending on the verb employed (e.g., *brát – bral* or *chtít – chtěl*).

4.3 Methodology

4.3.1 Ethics

The experiment was approved by the Charles University Ethics Committee since it met all the requirements on how the data were anonymized and how personal information was handled.

4.3.2 Recruitment

As in the previous experiment, the participants were recruited via an e-mail that included all the necessary information about the research. The second-language learners were recruited from the students of the Faculty language school whose English level had been tested by a placement test.[26] The use of dictionaries and other study materials was prohibited during the testing. Native speakers of English were recruited by an advertising e-mail sent to the selected students of the Anglo-American University in Prague.

Participants with a required language level and first language were addressed directly. A small fee was given to all participants (as this research project has been financially supported by a Specific Academic Research grant, "Jazyk a nástroje pro jeho zkoumání").

4.3.3 Participants

88 English second-language learners with Czech as their first language at A1 – C1 proficiency levels (16 at the A1 level[27], 15 at the A2 level, 18 at the B1 level, 19 at the B2 level, and 20 at the C1 level) and a control group of 9 native speakers[28] were recruited. None of the participants had been diagnosed with any language impairment. The mean age was 37 years, standard deviation 9.4 years. 73 participants were female, 24 were male.

[26] For more information on the placement test, see Chapter 3.

[27] According to the English Grammar Profile, the A1-level students already possess the ability to use regular and irregular main verbs in simple verb phrases. The grammar of the English past tense is also taught within the first lessons of any English beginner course (as evidence by the structure of English course books).

[28] A control group consisting of only 9 native speakers may not seem as a sufficient enough group for any comparison with the L2 speakers; however, our group of native speakers trully served the control purpose and our findings could be further compared to the previous work by Albright and Hayes (2003), who focused on adult native speakers of English, using the exact same task.

4.3.4 Consent

Similarly to the previous experiment, the testing took place in a quiet room located at the Faculty of Arts. The participants were asked to express their consent with the experiment, data anonymization, and research goals. They were also informed that all the data would be stored in a locked cabinet to which only the administrators would have access. Each participant was then given a random five-digit number[29] in order to anonymize their data.

4.3.5 Stimuli

The study adopted 32 novel verbs that were deemed phonotactically legal in English[30] (e.g., *bize*, *drice*, *spling*) from Albright and Hayes's (2003) and Blything et al.'s (2018) studies. Some of the original novel words were excluded due to their existence in Present-Day English (since the novel words were created in 2003 and some of them have already entered the language as real words since) and the need for a balanced dataset.

4.3.5.1 The creation of the stimuli

Albright and Hayes (2003) started with the creation of a control corpus of existing English verbs, which were all sourced from the English part of the CELEX lexical database (Baayen et al., 1995), with a special focus on these verbs whose lemma frequency equalled or exceeded 10. If the verb had more than one past-form variant (such as *dive-dived/dove*), both alternatives were included separately. The resulting corpus of existing English verbs consisted of 4,035 regular and 218 irregular verbs. Those verbs served as evidence of real phonotactic properties and phonological changes in past-tense forms and were used for counting the (dis)similarity of the created novel-verb stimuli to the existing English (ir)regular verb. In order to do that, the phonemic transcriptions of those verbs were fed into Albright and Hayes' rule-based model (Albright & Hayes, 2003).

Albright and Hayes used the same model for the creation of their novel verbs. They created a dataset of 2,344 novel-verb forms by "concatenating combinations of relatively common syllable onsets and syllable rhymes" (Albright & Hayes, 2003: 135). The full list of

[29] Generated at https://www.random.org/.

[30] Phonotactically legal English novel words follow the restrictions on the permissible combinations of English phonemes (e.g., Pitt, 1998).

novel-verb forms was then uploaded back into the model, which then produced (ir)regular past forms for each novel verb and also rated their similarity to existing (ir)regulars (Albright & Hayes, 2003). For each novel word, Albright and Hayes (2003) and Blything et al. (2018) employ two ways of counting the novel word's similarity to existing (ir)regulars – based on (i) the Generalized Context Model (hereby GCM; see Nosofsky, 1990) and (ii) the Minimal Generalization Learner (hereby MGL; see Albright & Hayes, 2003). GCM provides values of (i) "variegated" similarity, while MGL outputs (ii) "structured" similarity (Albright & Hayes, 2003; Blything et al., 2018).

4.3.5.2 "Variegated" similarity measure (GCM; Nosofsky, 1990)

Albright and Hayes (2003) implemented Nosofsky's (1990) Generalized Context Model (GCM), using a "variegated" similarity measure. GCM is programmed to output "a well-formedness score for a novel verb's past-tense form" (Blything et al., 2018: 6). The model first searches for the phonological change by which the novel verb generates its past form (for instance, [aɪ] → [o] for *scride* to become *scrode*). Then it finds all real English verbs with a similar past-tense change in phonology (e.g., *ride-rode*) and generates the score that tells us how similar the novel verb appears to all real verbs undergoing similar past-tense change (Blything et al., 2018). GCM compares each phonological part of the novel verb to the related real verb based on Broe's (1993) estimate of relative similarity (Albright & Hayes, 2003). These similarity scores of the novel verb to all related existing verbs are added up and output "an overall well-formedness score" (Blything et al., 2018: 6), i.e., how similar the new item is to real regular and irregular verbs.

4.3.5.3 "Structured" similarity measure (MGL; Albright & Hayes, 2003)

The alternative measure of phonological similarity used in Albright & Hayes (2003) based on a "structured" similarity measure was obtained from a different model, the so-called Minimal Generalization Learner (MGL). MGL is capable of generating "micro-rules" that describe phonological changes which a given verb has to undergo during past-tense formation and actively looks for phonological contexts in which those changes happen (this process is called "minimal generalization," see Blything et al., 2018). The model then outputs "graded confidence values for a verb's past-tense forms based on the reliability of the context-dependent micro-rule which produces that form" (Blything et al., 2018: 6). In comparison to CGM, this model requires not only for all comparison verbs to share the same phonological change when forming the past-tense forms (for instance *ø → əd* in verbs ending in [d] or [t] with the inflectional suffix *-ed*, e.g., *wan**t*** or *deci**de***) but they also need to share the same phonological properties (e.g., the verb must end in the alveolar [t] or [d]) (Blything et al., 2018).

4.3.5.4 Comparison and the final decision

Since Albright and Hayes (2003) and Blything et al. (2018) introduce two different models of similarity measures, it was necessary to decide which of the two models will be used in our data analysis. As Blything et al. (2018: 6) argue, GCM (Nosofsky, 1990) compares "phonological segments across entire words such that phonological similarities at the beginning, middle, and end of a word have equal weighting, and each comparison form can be similar to the novel verb in its own way." The stimulus' similarity to each related word is compared "feature-by-feature" (Ambridge, 2019: 12) (e.g., the change from *i* to *a* in *sit, spit,* or *swim* when forming past tense) and is then divided by its feature-by-feature comparison to all existing verbs stored in our mental lexicon. The result yields the probability of a given word being a part of a specific class. For that reason, the variegated similarity is sometimes regarded as a "pure" analogy by some linguists (Albright & Hayes, 2003: 122). On the contrary, MGL (Albright & Hayes, 2003), working with the structured similarity, pays attention "only to phonological properties that are shared uniformly among comparison verbs (e.g., the verb *cling* shares its final two segments with *string*, *sting*, and *fling,* etc.)" (Blything et al., 2018: 10) and thus assess the similarity to the phonological properties based on the features "most relevant to a verb's past-tense form (e.g., the past tense of *cling* is *clung* because it shares a final segment with *string*, *sting*, and *fling*)" (Blything et al., 2018: 10), creating explicit micro-rules for each sub-regularity (Ambridge, 2019). Measures of variegated similarity of *cling* may be thus misled by irrelevant segments of a word. Blything et al. (2018: 7) conclude that as such, "structured similarity is compatible with more sophisticated conceptions of analogy such as schema-based approaches." Considering possible shortcomings of the variegated similarity measure obtained from GCM, the analysis will use the structured similarity measure as a primary decisive metric.

Albright and Hayes (2003) then further divided the 32 novel words into four islands of reliability – 8 novel words resembling regulars (e.g., *blafe* or *tesh*), 8 novel words resembling irregulars (e.g., *chake* or *teep*), 8 novel words resembling both regulars and irregulars (e.g., *glip* or *stin*), and 8 novel words resembling neither regulars nor irregulars (e.g., *pank* or *rask*).

Albright (2002) defines these islands of reliability as "phonological contexts in which a particular morphological change works especially well in the existing lexicon as islands of reliability" (Albright & Hayes, 2003: 127), building on a hypothesis that "speakers know the contexts in which the regular change can be relied upon to a greater than average extent" (Albright & Hayes, 2003: 127). Since these islands are applicable to both regular and irregular verbs, Albright and Hayes (2003) have also included the island of reliability alongside the similarity to existing regulars and irregulars for each created novel word.

4.3.5.5 Cross-language interference effects

Since the thesis also attempts to explore the effect of L1 on second-language production, cross-language interference effects have been taken into consideration. The 32 novel verbs were further divided into 16 phonotactically legal (A set) and 16 phonotactically illegal (B set) novel words for Czech (for the full table of novel word stimuli, see Appendix 2).[31] Since there is no existing counter of the novel word's similarity to an existing Czech word, Šturm & Lukeš (2017)'s paper "Fonotaktická analýza obsahu slabik na okrajích českých slov v mluvené a psané řeči" (with a focus on the spoken part) was used as reference for the division. This study investigates Czech "syllable onsets and final codas" (Šturm & Lukeš, 2017: 99) both in written and spoken Czech texts imported from the Czech National Corpus. Along with a detailed study on the complexity of Czech syllable onsets and codas, the authors have also compiled a list of all phonotactically legal Czech syllable onsets and codas.[32] Using this reference list, and more specifically its spoken part (since the experiment was equipped with an aural prompt), our stimuli were further divided into 16 words that were phonotactically legal in Czech (i.e., both the syllable onset and coda were deemed phonotactically legal in Czech) and into 16 words phonotactically illegal in Czech (i.e., either the syllable onset or coda or both parts were deemed phonotactically illegal in Czech). One of the major downsides of this division is that Šturm and Lukeš's paper does not investigate syllable nuclei, and our stimuli division, therefore, could not provide a full and complex picture of Czech phonotactic (il)legality. However, given the lack of any automatized calculator that would allow stimuli importation and export the word's similarity or dissimilarity to existing Czech words, this phonotactic analysis was the only relevant reference point that could be used for the division. Therefore, even though we are fully

[31] Phonotactically legal Czech novel words follow the restrictions on the permissible combinations of Czech phonemes (e.g., Pitt, 1998 – an experimental study done on English; its findings are transferable to other languages, though, Czech included).
[32] Available at https://fonetika.ff.cuni.cz/vyzkum/materialy/fonotaktika/.

aware of its drawbacks, using this paper was, at the moment, the only suitable reference point for the inclusion of Czech interference effects available. Any assumptions on the effect of Czech, however, need to be read with this in mind.

The reason for including phonotactic legality in Czech in the experiment that otherwise involves L2 words was our intention to investigate whether participants will react differently to novel words phonotactically legal or illegal in Czech. We have hypothesized that if they do, Czech (the participants' L1) might be influencing the behaviour. An even more useful procedure would be involving a measure of a novel word's phonotactic probability in Czech; yet, as with the phonotactic legality, there is no automatized counter available for Czech at the moment (see also Conclusion).

4.3.5.6 Frame sentences

The novel words were then incorporated into frame sentences sourced from Blything et al. (2018). The original sentences needed to be adapted lexically (using the English Vocabulary Profile[33]) for the lowest-level participants, i.e., the A1 level, to ensure the study's applicability and participants' understanding. Each frame sequence contained four sentences presenting the novel word in a bare-infinitive, present-progressive, and present-simple form of the following template: "The (agent) likes to VERB. Look, there he is VERBing. Every day he VERBs. So yesterday he…" (e.g., *The baby likes to dize. Look, there he is dizing. Every day he dizes. So yesterday he…*). In order to avoid semantic effects of the surrounding words (as documented, for instance, in Ramscar (2002)) on the choice of past-tense form, each novel word was inserted into three different frame sentences (e.g., *This person/the waitress/my father likes to dize. Look, there s/he is dizing. Every day s/he dizes. So yesterday s/he…*; see Appendix 2 for the full list of the frame sentences). The frame sentences and their pairing with the novel verb were random and therefore different for every single participant.

[33] Accessible at https://www.englishprofile.org/wordlists/evp.

4.3.5.7 Stimuli recording

The stimuli were recorded in the sound booth of the Institute of Phonetics at the Faculty of Arts, Charles University, by a native speaker of English who had been previously instructed to record the frame sentences with proper intonation and a special focus on the novel word.

4.3.6 Procedure

In an elicited production paradigm, each participant was presented with the 32 novel verbs inserted into frame sentences. The participants were told they would hear sentences in which someone would perform an activity described with a novel word. They were then instructed to finish the sentence when the recording stopped playing. The participants were thus prompted to say what the agent "did yesterday," and past-tense forms elicitation was ensured.

The experiment was created in the psycholinguistic software PsychoPy v3.0 (Peirce et al., 2019). Each participant was first presented with two practice frames with a simplified visual transcription of the recording (e.g., "The (agent) likes to VERB. Look, there he is VERBing. Every day he VERBs. So yesterday he…" to avoid transcribing the novel word orthographically) to ensure that even the lowest language levels got accustomed to the experiment structure and understood the frame sentences. The trial session was followed by a test session without any visual support, focused solely on the aural input. All novel verbs were presented to the participant in random order and were, therefore, different for each participant. The recorded measures obtained from this experiment included the produced forms and reaction times to the presented novel word.

4.4 Results

4.4.1 Analysis of the produced forms

This analysis attempts to answer the question of whether the produced form is influenced by the novel word's similarity to existing regulars/irregulars or not. A total of 3,094 trials were recorded (508 for the A1 level, 477 for the A2 level, 574 for the B1 level, 608 for the B2 level, 640 for the C1 level, and 287 for the native speakers). Responses were coded using the response scale used by Blything et al. (2018), i.e., according to whether the recorded output form was (i) a regular past-tense form (e.g., *blafed*), (ii) an irregular past-tense form[34] (e.g., *blofe*), (iii) a form without any formal change (e.g., *blafe*), (iv) a third-person singular of present simple (e.g., *blafes*), (v) a progressive form (e.g., *blafing*), or (vi) a form unclassifiable under (i), (ii), (iii), (iv), or (v) (e.g., *blafest*). The mean proportion of each output form by the language level is given in the table below.

	Regular (vs. irregular only)	Regular (vs. all)	Irregular (vs. all)	No change (vs. all)	Third-person present (vs. all)	Past progressive (vs. all)	Unclassified (vs. all)
A1 level	94% (vs. 6%)	86%	5%	1.6%	0%	0.14%	6%
A2 level	91% (vs. 9%)	82%	8%	3%	0.6%	0%	7%
B1 level	87% (vs. 13%)	68%	12%	2%	0.3%	0%	2%
B2 level	83% (vs. 17%)	79%	16%	2%	0%	0.3%	2%
C1 level	82% (vs. 18%)	79%	17.5%	2%	0%	0.8%	1%
Native speakers	66% (vs. 34%)	65.5%	33%	0.7%	0%	0%	0.3%
All learners	87% (vs. 13%)	82%	12%	2%	0.2%	0.5%	3%
All participants	85% (vs. 15%)	80%	14%	2%	0.2%	0.5%	3%

Table 14. The descriptive statistics for the analysis of produced forms. The table was sourced from Blything et al. (2018).

For each trial, the produced forms (and their coding) were recorded (along with the ID number, the language level, the placement test score, the gender, the set, and the MGL scores). Since this study is focused on the development of the past-tense system, our analysis focused solely on the investigation of regular and irregular forms, and responses (iii) – (v) were excluded from the analysis.

To investigate whether the learner data showed similar outcomes to Albright & Hayes (2003) and Blything et al. (2018), the data were analysed with a generalized mixed-effects

[34] Contrary to the approach adapted in Blything et al. (2018), we have decided to include not only irregular forms of the novel verbs that "received the most favourable score in Albright and Hayes's (2003) study" (Blything et al., 2018: 16), but also all the other possible irregular past forms cited in Quirk et al. (1985: 104). As a consequence, the irregular past forms of verbs *drit, gude,* and *nold* were homophonous with the stem form based on its analogical similarity to no-change forms such as *cut, put*, or *hit.* Such past forms were then coded as "irregular" rather than "no change."

model[35] (using the *lme4* package; see Bates et al., 2015) in R (R Development Core Team, 2011). Linear mixed-effects models were again chosen as a powerful statistical method that works with individual trials, is robust against missing data and against abnormal data distribution, and enables us to treat participants and individual novel verbs as random effects (Baayen et al., 2008).

We decided to keep the structure of random effects in its simplest and most-frequently-used form of (1|part) + (1|item) with "part" standing for participants and "item" for individual novel words. The dependent variable used was the production of either regular or irregular past forms; the independent variables then varied according to the focus of each particular analysis and included (a) the level, (b) the similarity to existing regular verbs, and (c) the similarity to existing irregular verbs. All independent variables were imported into the model at the same time to enable us to investigate their interactions. To account for possible multicollinearity of the independent variables, the variance inflation factor (VIF) was checked for each model. The values for the initial analysis were as follows: level - 1; similarity to existing regulars – 1.05; similarity to existing irregulars – 1.05.

4.4.1.1 Initial data analysis

First, an analysis was run with the language level, the similarity to regulars, and the similarity to irregulars as independent variables, using the MGL probability model to see if participants at different levels performed differently (and, therefore, if group post-hoc tests were needed) and whether answers were predicted by the similarity to existing regulars or irregulars. Again, we followed Pérez et al. (2016)'s procedure of obtaining the most precise and explanatory models. We started the analysis with the highest number of possibly interacting independent variables: the level, similarity to existing regular verbs, and similarity to existing irregular verbs, and compared it with simpler interactions using a stepwise model comparison (see Pérez et al., 2016). Since the deletion of this complex three-way interaction gave a statistically significant outcome ($p < .001$), the three-way interaction was therefore kept in the model. All independent variables were imported into the model simultaneously to investigate their potential interaction. The overview of the generalized linear model can be seen in Appendix 2. To get a simpler overview of the significant variables, the analysis of variance (ANOVA) was then run on the model. It can be seen in the table below.

[35] The data meet all the necessary assumptions of the model.

	Df	Sum Sq	Mean Sq	F value	P-value
Level	5	25.692	5.138	**5.138**	**0.0003**
Similarity to regulars	1	2.993	2.993	*2.993*	*0.087*
Similarity to irregulars	1	0.923	0.923	0.923	0.339
Level:similarity to regulars	5	6.988	1.398	1.398	0.233
Level:similarity to irregulars	5	10.729	2.146	*2.146*	*0.067*
Similarity to regulars: similarity to irregulars	1	0.291	0.291	0.291	0.591
Level:similarity to regulars: similarity to irregulars	5	11.431	2.286	*2.286*	*0.053*

Table 15. The analysis of variance table for the advanced analysis of the produced forms. The numbers in bold indicate a significant effect, the numbers in italics indicate marginal significance.

The results show a significant effect of the language level, indicating that participants at different levels perform differently. Participants' performance could be further influenced by the similarity to regulars (i.e., the novel word's similarity to regulars predicts the type of the form used) since the p-value is still marginally significant. The novel word's similarity to existing irregulars then likely influences the performance in various language groups differently since the p-value is again marginally significant. A marginally significant effect of the interaction between the three independent variables also suggests that the interaction between the similarity to regulars and the similarity to irregulars is prone to change according to the language level. In order to offer a visualization of the three-way interaction between level, similarity to regulars, and similarity to irregulars, we will provide two charts: one plotting level and similarity to regulars, and another one plotting level and similarity to irregulars.

First, to get a general idea of how the language group performed in relation to the similarity to existing regulars and also to compare our findings with Blything et al. (2018), for each novel word at each language level, the mean proportion of regular inflections was counted and included in a scatter plot alongside the similarity to existing regulars. The figure below plots the mean proportion of regular inflection and its dependence on the novel verb's similarity to existing regulars for each language group separately, using a trendline to see the possibly significant development for individual groups.

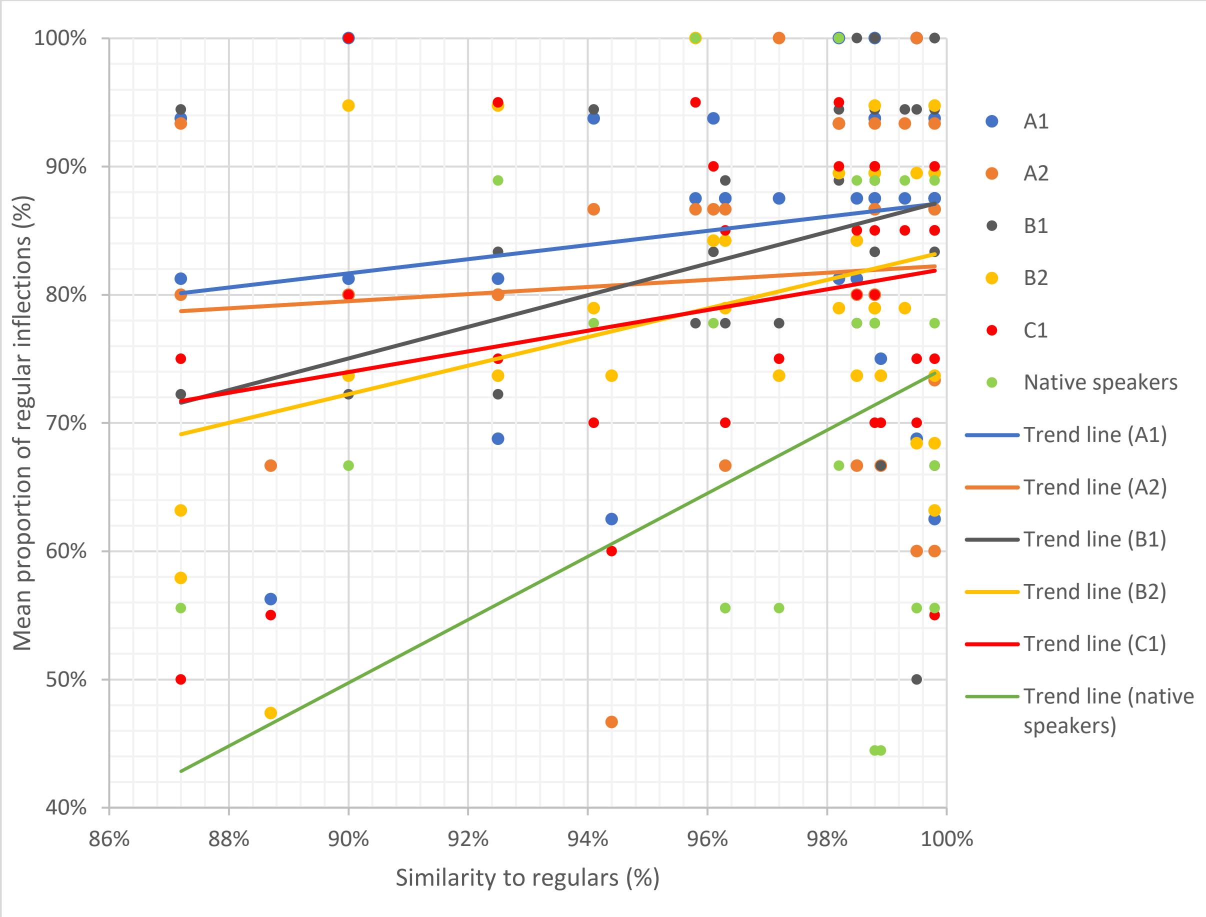

Figure 14. Mean proportion of past forms with regular inflection based on the novel verb's similarity to existing regulars, divided into language levels.

The plot shows that out of the language levels, A1 participants produced the highest number of regular past-forms, followed closely by the A2 level, B1 level, C1 level, B2 level (notice the order of the proficiency progression), and finally by the native speakers. The flattest trend line belongs to the A2 level: The A2-level learners seemingly are not sensitive to the novel word's similarity to regulars when producing the past-form, while the proportion of regular inflection stays relatively high. Other learner trendlines show more steepness, although the angle and consequently the awareness towards the similarity to existing regulars differ. The only group that seemed truly sensitive to the similarity to regulars is the control group consisting of native speakers – they produced less regular past-forms with novel words less similar to existing regulars while producing more regular past-forms with novel words more similar to existing regulars. The evident steepness of the trendline indicates real awareness towards the novel word's similarity to existing regulars. The L2 learners, on the contrary, tended to produce more regular past-forms in general.

The figure above suggests that the native speakers produce more regular past-tense forms with novel verbs that are phonologically highly similar to existing regular verbs, in line with the single-route model (and shown both in Albright & Hayes, 2003 and Blything et al., 2018). However, L2 learners seem to be less dependent on the verbal similarity to regulars when deciding on its past form and generally preferring the addition of the rule. In respect to how the similarity to existing regulars affects the production, the figure also indicates that more regular past-tense forms are generally produced with stimuli that are more similar to existing regulars, and fewer of them are produced the farther the stimuli move from existing irregulars.

Second, the effect of the interaction of the level and similarity to irregulars on the produced form (though only marginal in the initial ANOVA) has shown a similar development as in the preceding plot:

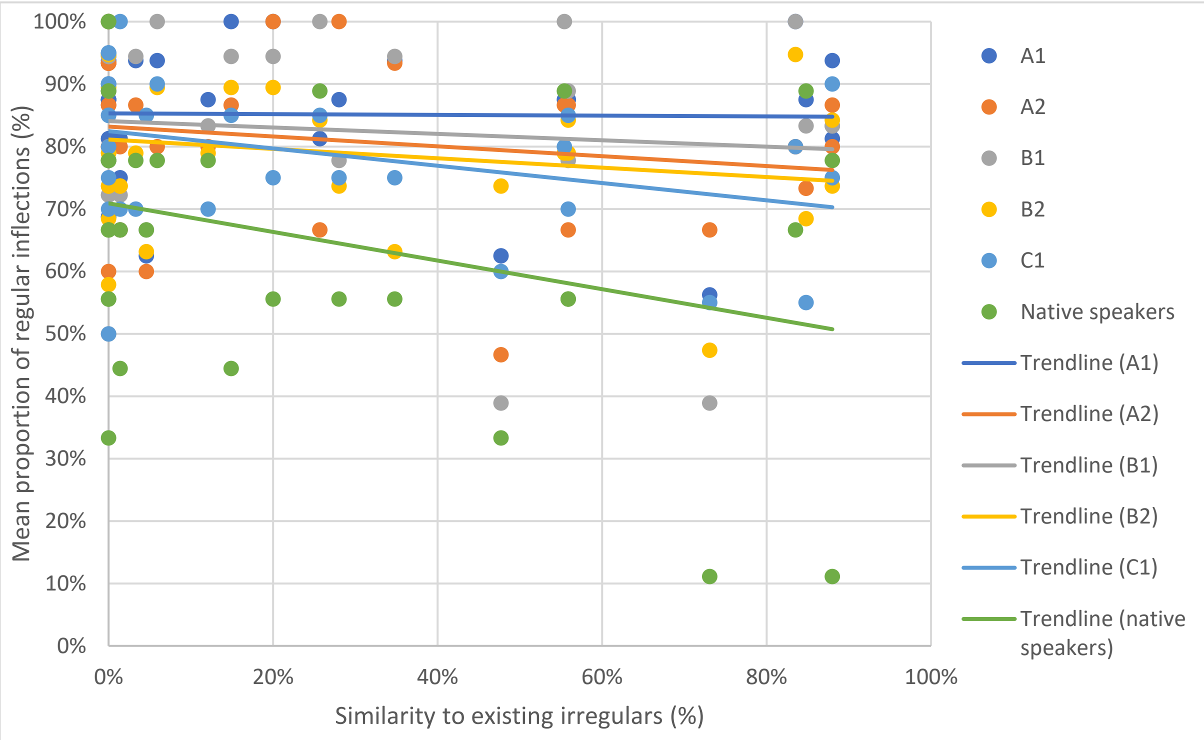

Figure 15. Mean proportion of past forms with regular inflection based on the novel verb's similarity to existing irregulars, divided into language levels.

Unsurprisingly, the mean production of regular inflections decreases with the novel word's similarity to existing irregulars. While the A1 level seems to be rather insensitive to such similarity due to the flatness of the trendline, the rest of the language levels show more steepness and consequently more awareness of the stimuli's similarity to existing irregulars. Out of the L2 proficiencies, it is the C1 level that seems to be the most sensitive to this similarity. The native speakers, similarly to the previous analysis, produced the least amount of regular inflections, and the steepness of their trendline also indicates real awareness of the novel word's similarity to existing irregulars.

The output of the ANOVA that was subsequently run on the linear mixed model, as summarized in the table above, has shown a marginally significant effect of the interaction between the level, similarity to regulars, and similarity to irregulars. To understand this performance fully, group post-hoc tests for each language level were needed. For each language group, a data subset has been created, and a generalized mixed model has been run, followed by an ANOVA, with the similarity to existing (ir)regulars as independent variables.

4.4.1.3 Group post-hoc tests

4.4.1.3.1 The A1 level

For the A1 level, the analysis of variance showed the following results:

	Df	Sum Sq	Mean Sq	F value	P-value
Similarity to regulars	1	0.146	0.146	0.146	0.703
Similarity to irregulars	1	1.046	1.046	1.046	0.309
The similarity to existing regulars:existing irregulars	1	0.559	0.558	0.559	0.457

Table 16. The analysis of variance table for the A1 subset.

As the results indicate, the performance of the A1-level participants was not predicted by either the similarity to existing regulars or existing irregulars. Such findings do not seem to correspond to the single-route model and incline more to a dual-route model of morphological productivity.

4.4.1.3.2 The A2 level

For the A2 level, the analysis of variance showed the following results:

	Df	Sum Sq	Mean Sq	F value	P-value
Similarity to regulars	1	0.325	0.325	0.325	0.570
Similarity to irregulars	1	0.003	0.003	0.003	0.756
The similarity to existing regulars:existing irregulars	1	0.082	0.082	0.082	0.775

Table 17. The analysis of variance table for the A2 subset.

Similarly to the A1 subset, the results indicate that the performance of the A2-level learners was not predicted by either the similarity to existing regulars or existing irregulars, corresponding more to a dual-route model of morphological productivity.

4.4.1.3.3 The B1 level

For the B1 level, the analysis of variance showed the following results:

	Df	Sum Sq	Mean Sq	F value	P-value
Similarity to regulars	1	0.700	0.700	0.700	0.405
Similarity to irregulars	1	0.427	0.427	0.427	0.515
The similarity to existing regulars:existing irregulars	1	0.103	0.103	0.103	0.749

Table 18. The analysis of variance table for the B1 subset.

As with the A1- and A2-level learners, the B1 subset's performance was not predicted by either the similarity to existing regulars or existing irregulars, corresponding more to a dual-route model of morphological productivity.

4.4.1.3.4 The B2 level

For the B2 level, the analysis of variance showed the following results:

	Df	Sum Sq	Mean Sq	F value	P-value
Similarity to regulars	1	4.615	4.615	**4.615**	**0.034**
Similarity to irregulars	1	1.030	1.030	1.030	0.313
The similarity to existing regulars:existing irregulars	1	0.917	0.917	0.917	0.341

Table 19. The analysis of variance table for the B2 subset. The numbers in bold indicate a significant effect.

In contrast to the previously analysed A1, A2, and B1 subsets, and more in line with the single-route model, a significant effect of similarity to existing regular verbs was found in the B2-level participants. This finding suggests that the novel verb's similarity to existing regulars is a decisive factor in the production of past-tense forms.

4.4.1.3.5 The C1 level

For the C1 level, the analysis of variance showed the following results:

	Df	Sum Sq	Mean Sq	F value	P-value
Similarity to regulars	1	1.422	1.422	1.422	0.236
Similarity to irregulars	1	3.307	3.307	*3.307*	*0.072*
The similarity to existing regulars:existing irregulars	1	1.441	1.441	1.441	0.233

Table 20. The analysis of variance table for the C1 subset. The numbers in italics indicate a marginally significant effect.

The C1 subset showed a marginally significant effect of similarity to existing irregular verbs on the produced form. This would suggest that the novel verb's similarity to existing irregular verbs is a decisive factor in the production of past-tense forms. Plotting was needed in order to understand the direction of this relation:

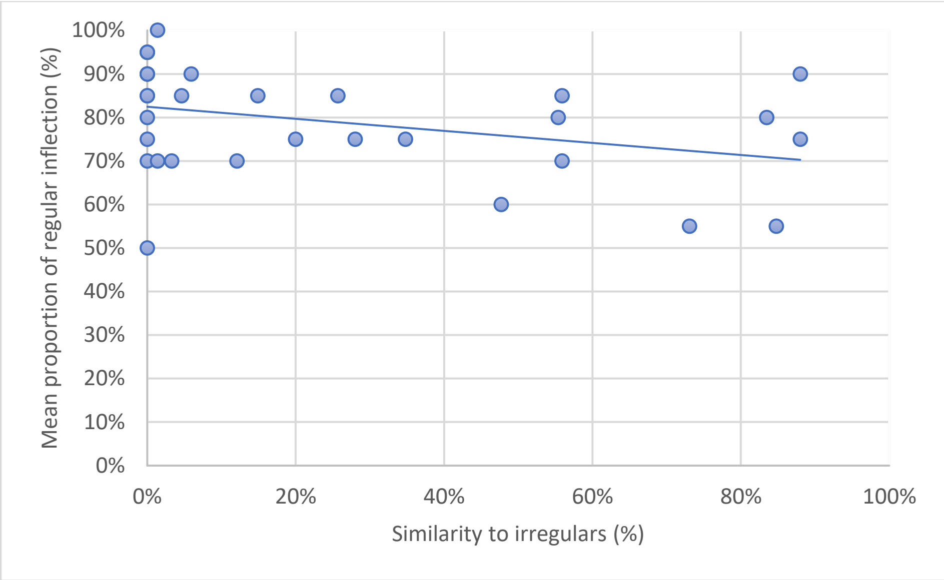

Figure 16. Mean proportion of past forms with regular inflection based on the novel verb's similarity to existing irregulars, as produced by the C1-level participants.

As the graphical visualization shows, the trend line of the mean production of regular inflections declines with the similarity to irregulars, which suggests that the C1-level learners produce less regular past-forms with the novel word that are more similar to existing irregulars.

4.4.1.3.6 The native speakers

For the native speakers, the analysis of variance showed the following results:

	Df	Sum Sq	Mean Sq	F value	P-value
Similarity to regulars	1	6.738	6.738	**6.738**	**0.011**
Similarity to irregulars	1	1.377	1.377	1.377	0.243
Similarity to existing regulars:existing irregulars	1	3.338	3.338	*3.338*	*0.071*

Table 21. The analysis of variance table for the native-speaker subset. The numbers in bold indicate a significant effect, the numbers in italics indicate marginal significance.

Similarly to the B2 subset, and in line with the single-route model, a significant effect of similarity to existing regular verbs was observed with the native speakers, implying that the novel verb's similarity to existing regular verbs once again played an essential role in the production of past-tense forms. Moreover, the data show a marginally significant effect of the interaction between the similarity to existing regulars and irregulars.

To look closely at the interaction, two data subsets were created – the novel words were divided into (a) 13 novel words that were more similar to existing regular verbs and (b) 13 novel words less similar to existing regular verbs – and both subsets were plotted against the similarity to existing irregulars:

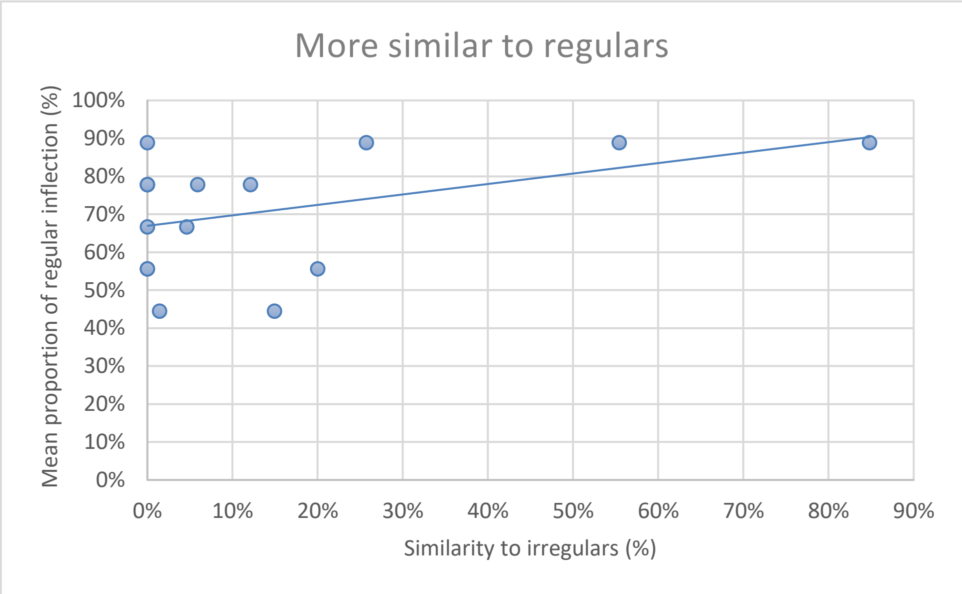

Figure 17. Mean proportion of past forms with regular inflections based on the novel verb's higher similarity to existing irregulars, as produced by the native speakers.

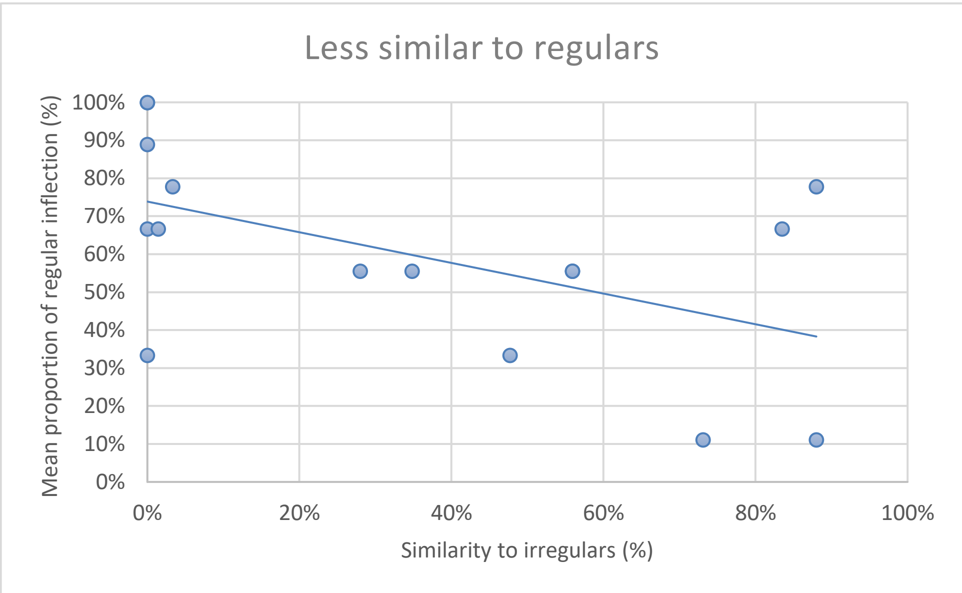

Figure 18. Mean proportion of past forms with regular inflections based on the novel verb's lesser similarity to existing irregulars, as produced by the native speakers.

The second figure shows an expected development - the more similar the novel word is to an existing irregular, the less likely is it produced as a regular past-form. However, the first figure shows an interesting trend: In the subset of novel words more similar to regulars, the more similar a novel word is to an irregular, the more likely are the items produced as regular past-forms. Given that the novel words in the subset are generally more similar to existing regulars, the participant may have a tendency to regularize (even the novel words that show the biggest similarity to existing irregulars out of the most regular ones).

4.4.1.4 Interim summary

So far, the initial data analysis of produced forms has shown that participants at different language levels perform differently, that participants' performance is likely influenced by the similarity to regulars, that the novel word's similarity to existing irregulars probably influences the performance in various language groups differently, and that the interaction between the similarity to existing regular verbs and the similarity to existing irregular verbs is prone to changes according to the language level. The graphical visualization of the data has shown that for the native speakers, the novel verb's similarity to existing regular verbs played a significant role in the production of past-tense forms, in line with the single-route model. However, L2 learners have shown lesser dependency on verbal similarity to regulars when deciding on its past form.

Group post-hoc tests have shown that the performance of the A1-, A2-, and B1-level participants was not predicted by either the similarity to existing regulars or existing irregulars. Such findings do not seem to correspond to the single-route model and incline more to a dual-route model of morphological productivity. In contrast to that, a significant effect of similarity to existing regular verbs was observed with the B2-level participants and native speakers. The C1-level participants have shown a marginally significant effect of similarity to existing irregular verbs, and the graphic visualization of the data has shown that they produce less regular past-forms with novel words that are more similar to existing irregulars. The native-speaker dataset has also indicated a marginally significant effect of the interaction between the similarity to existing regulars and irregulars, which, upon further graphical visualization, pointed at their tendency to regularize the novel words that are more similar to existing regulars and also to produce less regular past-forms with novel words more similar to existing irregulars.

Our findings suggest that the way by which L2 learners acquire the English past tense is characterized by the development from the mastery of mechanistic rules (at the A1, A2, and B1 levels) to the refinement of their application when the learner starts spotting analogical patterns of existing verbs (at the B2 and C1 levels, similarly to native speakers). In the development of inflectional morphology, the performance of second-language speakers thus comes closer to native speakers with the higher proficiency of the B2 and C1 levels, while the lower-proficiency learners seem to be using an alternative strategy of rule application.

4.4.2 Reaction-times analysis

The reaction-times analysis is designed to answer the question of whether reaction times, i.e., the length of time participants took to respond to a given stimulus, are different for novel words phonotactically legal and illegal in Czech. This would offer us a closer look into possible transfer effects at play during language production.

4.4.2.1 Editing the data set (removing outliers and superfluous forms)

A total of 3,094 reaction times were recorded (508 for the A1 level, 477 for the A2 level, 574 for the B1 level, 608 for the B2 level, 640 for the C1 level, and 287 for the native speakers). The trials were first checked for outliers (for our definition of an outlier, see Chapter 3). For the presented data, the first quartile starts at 0.344 seconds, and the third quartile ends at 0.864 seconds, thus creating an interquartile range of 0.52 seconds. The upper bound for the regular data equals 1.5*(TQ + IQR), i.e., 1.644 seconds. The lower bound then equals 1.5*(FQ – IQR), i.e., -0.436 seconds. All data points that do not fit into the range of -0.436 and 1.644 seconds were then considered outliers and deleted from the final data set. Since our primary interest was the difference between the reaction times of produced regular and irregular past-forms, the reaction times to other coded responses (i.e., no change or past progressive) were excluded from the dataset as well. After the omission of the outliers and superfluous responses, the dataset had the following characteristics:

	N	Mean reaction times (s)	Minimum reaction time (s)	Maximum reaction time (s)	Standard deviation (s)	Mean reaction times of regular forms (s)	Mean reaction times of irregular forms (s)
A1 level	440	0.617	0.06	1.64	0.344	0.614	0.665
A2 level	401	0.594	0.12	1.55	0.311	0.593	0.598
B1 level	502	0.566	0.06	1.64	0.345	0.552	0.682
B2 level	532	0.530	0.08	1.64	0.306	0.520	0.589
C1 level	569	0.569	0.12	1.58	0.302	0.565	0.591
Natives	263	0.597	0.13	1.63	0.341	0.633	0.526
All learners	2,444	0.573	0.06	1.64	0.322	0.567	0.615
All particip.	2,707	0.575	0.06	1.64	0.324	0.572	0.594

Table 22. Descriptive statistics of the dataset used for the reaction-times analysis.

Similarly to the previous analysis, for each trial, reaction times were recorded (along with the ID number, the language level, the placement test score, the gender, the set, and the MGL scores). As mentioned above, only reaction times to regular and irregular forms were used in the analysis, omitting all other produced forms under the category iii to v.

Given the size of the analysed sample that consists of 97 participants, the analysis assumes normal distribution based on the central limit theorem (see, for instance, Kwak & Kim, 2017; or above in Chapter 3) and, therefore, employs parametric tests. To investigate whether the data showed significant effects of various independent variables on reaction times, the results were analysed with a linear mixed-effects model.[36] The dependent variable used was the reaction times to the presented stimuli; the independent variables varied based on the type of analysis and included (a) the level, (b) the produced form, (c) the set (i.e., the novel word's il/legality in Czech)[37], (d) the similarity to existing regular verbs, and (e) the similarity to existing irregular verbs. Independent variables were always imported into the model at the same time to investigate their potential interactions. To account for possible multicollinearity of the independent variables, the variance inflation factor (VIF) was checked for each model. The values for the reaction-times analysis were as follows: level - 1.02; similarity to regulars -1.09; similarity to irregulars - 1.1; set - 1; produced form - 1.01. The model used participants and stimuli items (novel verbs) as random factors (the simplest and most-frequently-used form of random effects as described above in the analysis of produced forms – see section 4.4.1).

4.4.2.2 Initial transfer analysis

Again, we followed Pérez et al. (2016)'s procedure of obtaining the most precise and explanatory models. Using the random effects for participant and item (see (1|part) + (1|item) above), we started the analysis with the highest number of possibly interacting independent variables: the level, the set, and the produced form, and compared it with simpler interactions, using a stepwise model comparison (see Pérez et al., 2016). Since the deletion of this complex three-way interaction did not give a statistically significant outcome ($p = 0.71$), it was removed from the model. Following the same procedure, other interactions were removed from the model as well: the interaction between the level and the produced form ($p = 0.67$) and between

[36] The data meet all the necessary assumptions of the model.

[37] The A set consists of novel words phonotactically legal in Czech, while the B set consists of novel words phonotactically illegal in Czech.

the set and the produced form (p = 0.92), keeping only the interaction between the level and the set (p = 0.04). Since the removal of the produced form from the model led to a marginally significant result of the ANOVA (p = 0.076), the final model included the level, set, and produced forms as independent variables and participants and items (novel verbs) as random effects: level:set + level + set + produced_form + (1|part) + (1|item). P-values were generated by the *lmerTest* package (Kuznetsova et al., 2017). First, an analysis was run on the whole dataset to see if participants at different levels performed differently (and, therefore, if group post-hoc tests were needed) and whether novel words in different sets led to different reaction times. The overview of the linear model can be seen in Appendix 2. To get a simpler overview of the significant variables, the analysis of variance (ANOVA) was then run on the model. It showed a significant p-value for the interaction between the language level and the set and a marginal significance of the produced form.

	Df	Sum Sq	Mean Sq	F value	P-value
Level	5	0.207	0.041	0.563	0.728
Set	1	0.039	0.039	0.531	0.472
Produced form	1	0.232	0.232	3.149	*0.076 .*
Level:set	5	0.849	0.170	2.310	**0.042** *

Table 23. The analysis of variance (ANOVA) table for the initial interference analysis.

These results suggest that participants at different language levels reacted differently to the novel words of the two datasets (i.e., set A and set B based on the novel word's phonotactic (il)legality in Czech), which hints at possible interference effects at certain language levels. The results also suggest that the reaction times of individual language levels are roughly similar and that they are marginally influenced by the produced form, i.e., that the amount of time needed for the production of regular and irregular past-forms is slightly different. This provides us with enough evidence to perform post-hoc tests and see how the results differ at each language level.

For each set (A or B) at each language level, mean reaction times were counted and included in a bar plot. The figure below then shows how the language levels differ in their reaction to novel words of the two sets (bars for each language level are color-coded).

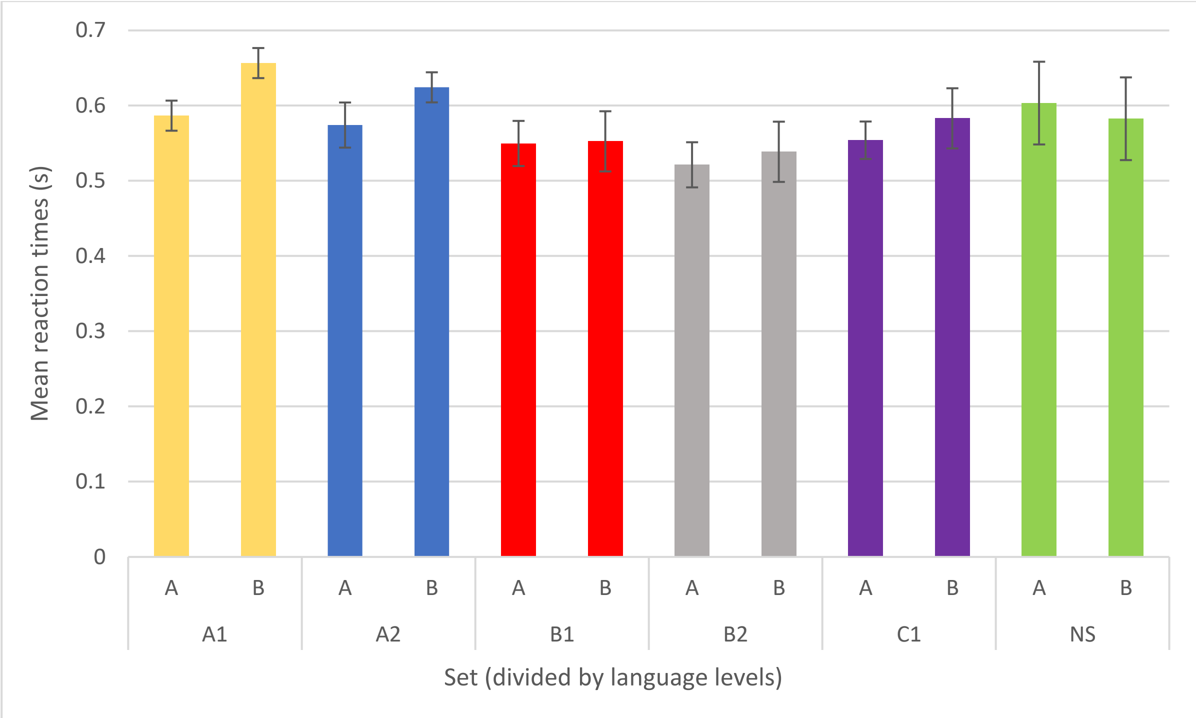

Figure 19. Mean reaction times to novel words of the two sets divided by language levels.

The plot shows that the L2 learners took longer to produce past forms of novel words from the B set, i.e., novel words that are phonotactically illegal in Czech, while the native speakers took longer to produce the A-set novel words, i.e., novel words phonotactically legal in Czech. These results would suggest that in the case of L2 learners, Czech has a facilitatory effect, i.e., it helps the participants react more quickly.[38] To see if the set actually had a significant effect on reaction times at each language level and if the difference in the mean reaction times between the A and B set was significant, a data subset was created for each language group, and an independent samples t-test was run between the data from the two sets.

4.4.2.1.1 Group post-hoc proficiency tests

For the A1 level, a significant difference appeared between the two sets $t(15) = -1.95$, $p = .035$, one-tailed, and a marginally significant effect when two-tailed, $t(15) = -1.95$, $p = .0695$. The results thus indicate that the reaction times at the A1 levels were predicted by the novel word's assignment to the set of novel words either phonotactically legal or illegal in Czech. The figure above also shows that the participants took longer with novel words

[38] Also notice how the quickness of reaction to presented stimuli rises with proficiency in L2 learners, except for the difference between the B2 and the C1 level, which has already been observed in Experiment 1 (Chapter 3).

phonotactically illegal in Czech, which suggests that at the A1 level, Czech acts as a facilitator, speeding up the reaction time to presented stimuli. In contrast, novel words from the B set, which are phonotactically illegal in Czech and thus lacking the resemblance to Czech words, showed longer reaction times.

For the A2 level, a marginally significant difference appeared between the two sets, t (14) = -1.69, p = .0562, one-tailed. Even though the results were only marginally significant, they indicate that the performance of the A2-level learners might still have been predicted by the set assignment of the novel words, similarly to the A1-level learners. The effect, though, does not appear to be as strong as with the A1 learners, and the facilitatory effect of Czech on reaction times will thus be likely weakened as well.

None of the B1, B2, or C1 levels showed a significant effect of the set on the reaction times or significant differences between the sets. For the B1 level, t (17) = -0.058, p = .95, two-tailed. For the B2 level, t (18) = -0.826, p = .42, two-tailed. And for the C1 level, t (19) = -0.807, p = .43, two-tailed. Those results suggest that from the B1 level to the C1 level, the participants' performance is not predicted by the novel word's assignment to one of the two sets. This indicates that no transfer effects are happening in response to the stimuli. Czech thus functions neither as the facilitator nor the inhibitor – the participants simply show no significant difference between the two languages at this stage of proficiency.

The figure above has also shown that, contrary to the L2 learners, the native speakers took longer to produce novel words phonotactically legal in Czech, which would suggest an inhibitory effect of Czech on the reaction times to the presented stimuli. However, the independent samples t-test did not show a significant difference between the two sets, t (8) = 0.501, p = .62, two-tailed. Those results suggest that similarly to the B1-, B2-, and C1-level participants, the native speakers did not significantly distinguish between the two languages during the past-form production, and Czech thus had no effect on their performance. Those four language groups made no difference between the novel words in the two sets and consequently between the two languages: This suggests that there were no transfer effects at play, and Czech functioned neither as a facilitator nor as an inhibitor.

4.4.2.1.2 The effect of the produced form

The initial analysis of variance has shown a marginally significant effect of the produced regular or irregular form on reaction times. The figure below shows the quickness of reaction to the two produced forms:

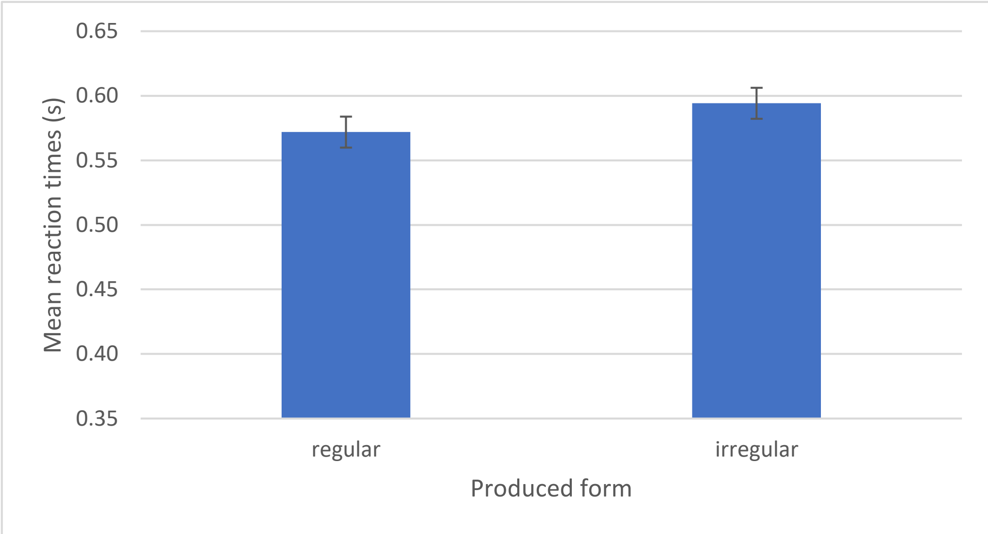

Figure 20. Mean reaction times to novel words of the two sets divided by language levels.

As to be expected, the graph above shows that producing irregular past forms of the stimuli took slightly longer for the participants, likely due to the application of analogy that appears to be more challenging for the participants.

4.4.2.3 Analysis of the similarities to existing (ir)regulars

Another question arises of whether the reaction times could be influenced by the novel word's similarity to existing regular or irregular past-tense forms. Again, we followed Pérez et al. (2016)'s procedure of obtaining the most precise and explanatory models. Using the random effects for participant and item (see (1|part) + (1|item) above), we started the analysis with the highest number of possibly interacting independent variables: the level, the similarity to existing regulars, and the similarity to existing irregulars, and compared it with simpler interactions, using a stepwise model comparison (see Pérez et al., 2016). Since the deletion of the three-way interaction did not give a statistically significant outcome ($p = 0.997$), it was removed from the model. Following the same procedure, other interactions were removed from the model as well: the interaction between the level and the similarity to regulars ($p = 0.78$) and between the level and the similarity to irregulars ($p = 0.72$), keeping only the interaction between the similarity to regulars and the similarity to irregulars (p = 0.04). The final model, therefore, included the following code of the similarity to regulars and irregulars independent variables and participants and items (i.e., the novel words) as random effects: MGL_R:MGL_IR + MGL_R + MGL_IR + (1|part) + (1|item). P-values were generated by the *lmerTest* package (Kuznetsova et al., 2017). The overview of the linear model can be seen in Appendix 2. To get a simpler overview of the significant variables, the analysis of variance (ANOVA) was then run on the model. It showed a significant p-value for the interaction between the similarity to existing regular verbs and similarity to existing irregular verbs and a marginal significance of the similarity to regulars.

	Df	Sum Sq	Mean Sq	F value	P-value
Similarity to regulars	1	0.304	0.304	4.100	*0.051* .
Similarity to irregulars	1	0.0001	0.0001	0.002	0.963
S. t. regulars:irregulars	1	0.336	0.336	4.542	**0.041 ***

Table 24. The analysis of variance (ANOVA) table for the advanced analysis of the isles of reliability. The numbers in bold indicate a significant effect, the numbers in italics indicate a marginally significant effect. Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1.

Those results suggest that the participants reacted differently to the novel words with varying similarities to existing regular and irregular verbs. Given that the deletion of the level from the initial code turned out insignificant, the results also suggest that the reaction times of individual language levels to those novel words are similar. Our findings also suggest that the quickness of reaction is slightly dependent on the novel word’s similarity to existing regulars. Graphical visualizations are needed to understand the interaction between similarity to regulars and irregulars in greater detail.

First, items were divided into four groups, and mean reaction times were counted for each group: (a) novel words less similar to regulars; (b) novel words more similar to regulars; (c) novel words less similar to irregulars, and (d) novel words more similar to irregulars (based on the division of MGL values into two halves for both regulars and irregulars). Mean reaction times for each group were then included in a bar plot. The figure shows how the participants differ in their reactions to the four groups (bars for each similarity group are color-coded).

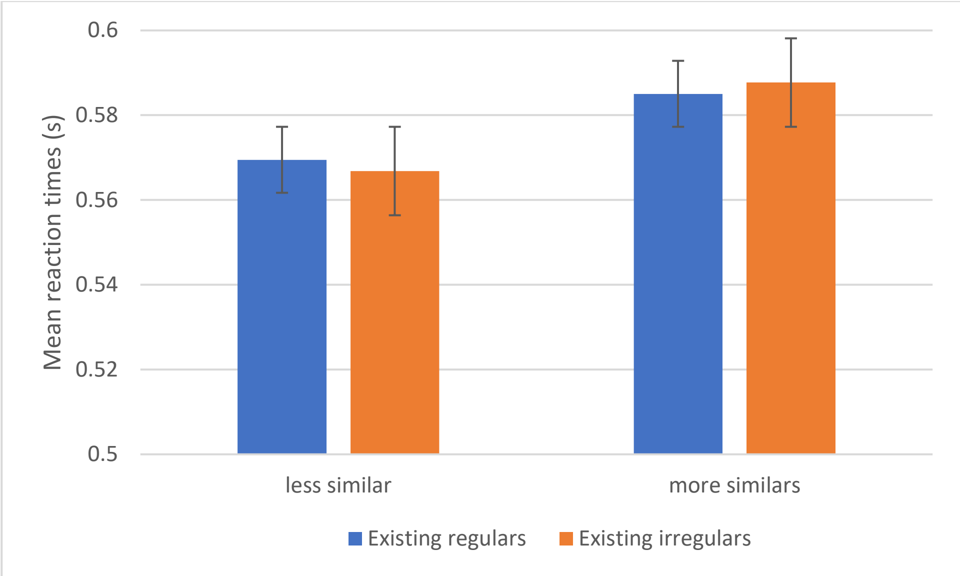

Figure 21. Mean reaction times to novel words that are less/more similar to existing regulars and irregulars.

The plot shows that in the group less similar to given existing verbs, the reaction to existing irregulars takes less time than the reaction to existing regulars. Novel verbs less similar to existing regulars, therefore, take more time to process (and consequently have their past-tense produced) than novel verbs less similar to existing irregulars. On the contrary, in the group more similar to existing verbs, the reaction to existing irregulars takes longer than the reaction to existing regulars. Novel verbs more similar to existing irregulars, therefore, take more time to process than novel verbs more similar to existing regulars. In general, producing a past-tense form of a novel word highly similar to existing irregulars takes longer, most likely due to its higher demand on human cognition, just as the production of a past-form of a novel verb less similar to regulars takes longer, most likely due to its unfamiliarity for the L2 learner.

4.4.2.4 Summary of the findings

The data analysis of the reaction times to presented past-tense sentences has shown that participants at different language levels react differently to novel words assigned to one of the two sets. The graphical visualization of the data has shown that the L2 learners took longer with novel words phonotactically illegal in Czech, while the native speakers took more time with novel words phonotactically legal in Czech. To investigate the statistical significance of the facilitatory or inhibitory effects of Czech, group post-hoc tests were conducted. The post-hoc tests showed that the A1-level learners, alongside the A2-level learners (though with a weaker effect), took significantly less amount of time with words phonotactically legal in

Czech, thus indicating that Czech functions as a facilitator in their performance. The B1, B2, and C1 levels then performed similarly to the native speakers, showing no significant effect of the set on the reaction times. Those four language groups thus made no difference between the novel words in the two sets and consequently between the two languages: This suggests that there were no transfer effects at play, and Czech functioned neither as a facilitator nor as an inhibitor. The data analysis and a further graphical visualization have also shown a marginal significance of the produced form on reaction times, with participants taking longer to produce an irregular past form.

The analysis of similarities to regular and irregulars and their effect on reaction times has shown that the participants took more time with novel words less similar to existing regulars and also with novel words more similar to existing irregulars, suggesting that the production a past-tense form of a novel word highly similar to existing irregulars causes higher demand on human cognition, just as the production of a past-form of a novel verb less similar to regulars indicates unfamiliarity of the L2 learner with the form.

4.4.3 Qualitative analysis

The last analysis of the data collected in the second experiment allows us to look qualitatively at a few interesting features of the data and their changes over the process of second-language acquisition. Attention will be paid to (i) different past forms produced by the participants, their mean proportion, and distribution over language levels; (ii) the pronunciation of the bound past-tense morpheme and its changes during the L2 development; (iii) the addition of superfluous morphemes to the past-tense endings; (iv) the additional lexicalization of novel words during the production of the past forms; and we will conclude with an overarching (v) item analysis.

4.4.3.1 The past-form analysis

Since this study is focused on the development of the past-tense system, it also raises the question of which forms are produced more frequently at each language level, if this production changes over the course of L2 acquisition, and if it is fuelled by any specific strategies. For those purposes, the produced regular and irregular past forms were subjected to qualitative analysis (along with those no-change forms, some of which were coded as irregular using Blything et al.'s (2018) response scale (i.e., *drit, gude, nold*)). The qualitative analysis first addresses the mean proportion of various past forms, then moves onto the types of irregular past forms, and concludes with the analysis of no-change past forms.

4.4.3.1.1 Mean proportion of various past forms

To investigate the proportion of regular and irregular forms and whether this proportion changes over the process of second-language acquisition, means were calculated for each past form at each language level. The distribution is visualized in the figure below:

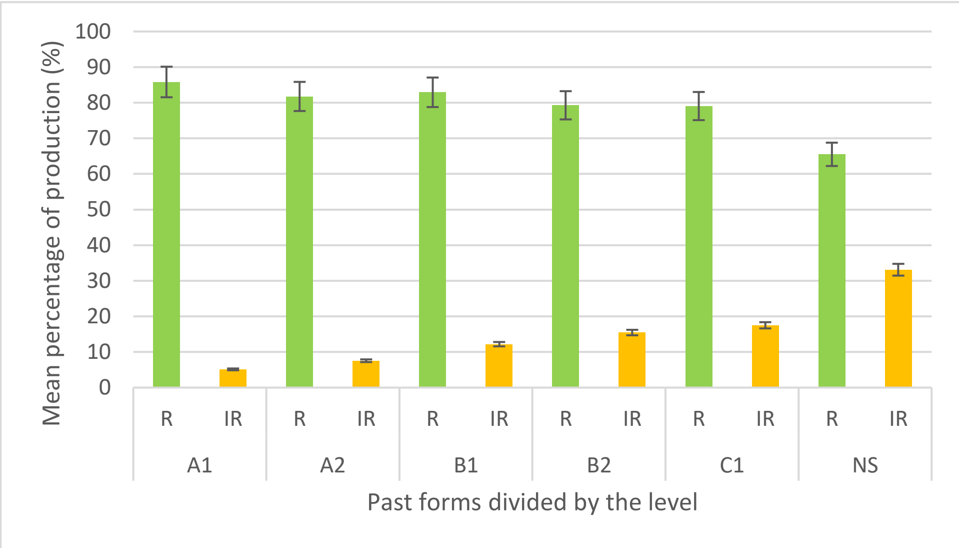

Figure 15. Mean percentage of production as a function of produced past forms divided into language levels (R refers to a regular past form, IR refers to an irregular past form).

The graphical visualization shows that at all language levels, the production of regular forms highly surpasses the production of irregulars. The plot also indicates that the production of regular forms declines with proficiency and that, in contrast, the production of irregular forms rises with higher proficiency. The mean production of regular past forms appears to be higher in L2 learners than in native speakers, and similarly, the mean production of irregular past forms seems higher in native speakers than in L2 learners.

4.4.3.1.2 Types of irregular past forms

Since we reported a higher proportion of irregular past-form production in more proficient participants, it is interesting to look more deeply into the types of irregular past forms used by the participants and whether their distribution changes with proficiency. Three occurring types of irregular past forms have been identified:

1. an internal change (exemplified by the change of /eɪ/ into /əʊ/ in *to chake – choke*)
2. an internal change accompanied by the addition of the past bound morpheme *-d, -t*, or *-ed* (exemplified by *to teep – tept* or *to bredge – bridged*)
3. an irregular no-change form (exemplified by *to drit – drit* or *to gude – gude*)

Further analysis then investigated the distribution of the three types over the language levels and its possible progression. The distribution is described in the table below:

Language level	Internal change + *-ed*	Irregular no change	Internal change
A1 level	61.54 %	34.62 %	3.85 %
A2 level	55.56 %	27.78 %	19.44 %
B1 level	51.43 %	20.00 %	28.57 %
B2 level	37.23 %	8.51 %	54.26 %
C1 level	30.36 %	9.82 %	59.82 %
Native speakers	37.89 %	7.37 %	54.74 %

Table 25. Mean percentage of production as a function of types of past forms divided by the language level.

The table above shows that the internal change accompanied by the addition of a past bound morpheme *-ed* seems to be a dominant production pattern at the lower levels of L2 learners, and its use slowly declines with higher proficiency, up to the point where it is predominantly, substituted by other production forms. Similar instances of this type have already been mentioned in the literature. Bybee and Slobin (1982), for example, suggested that the final *-t/d* may serve as an aid in acquiring the vowel change, i.e., as an important clue that the form is, in fact, expressing the past tense. In the earliest stages, children (and, in our case, the lower-level L2 students) have the knowledge of the regular *-ed* rule. Using it does not force them to rely so heavily on the surrounding context (as in the case of a sole internal past-tense change, for instance, in *sing-sang* or *break-broke*) to distinguish the past tense. The production of the past form both with the internal change and the *-t/d* suffixation may thus be understood as an intermediate mastery of the internal-change production (Bybee & Slobin, 1982). The L2 speakers may also be influenced by such models as *sleep/slept* or *keep/kept* (Rumelhart & McClelland, 1986). Rumelhart and McClelland (1986) have also noted that three of their four responses with this type of past-tense change were verbs ending in /p/, similarly to *sleep/keep*. They understand this as the participants' "sensitivity to the regular and subregular patterns of the English past tense" (Rumelhart & McClelland, 1986: 34).

At the lowest levels, this production is closely followed by irregular no-change forms; this strategy, too, shows a steady decline with higher proficiency. On the contrary, the third production type, the internal change, is in the minority at the lowest levels and gathers noticeable strength with higher proficiency. The native speakers, then, seem to be closer to the B2 and C1 levels in terms of the past-form distribution. Similarly to those two levels, the native speakers produce mostly internal-change past forms, followed by the internal changes with the bound morpheme and the internal change, respectively.

4.4.3.1.3 No-change past forms

In the previous section, we have investigated the changes in the distribution of irregular no-change past forms. An interesting question arises whether there can be a pattern for the participants to choose the no-change form as a past form and whether this pattern can be phonologically conditioned. When we look not only at irregular no-change past forms, i.e., *drit, gude,* and *nold,* but also at other no-change past forms such as *skride* or *gleed*,[39] there seems to be a phonologically-driven pattern in the no-change production. These five no-change forms appear relatively often in the participants' responses, seemingly regardless of the language level (along with other no-change forms that, however, do not appear consistently). The tables below describe the overall production of individual novel words within the no-change paradigm:

[39] For the novel-verb forms that were given the most favourable irregular score, see Albright and Hayes (2003). The irregular past forms of verbs *drit, gude,* and *nold* were homophonous with the stem form based on analogy with such no-change forms as *cut, put*, or *hit.* As stated above, such past forms of those three novel words were then coded as "irregular" rather than "no change."

Language level	*drit*	*gude*	*nold*
A1 level	55.56 %	22.22 %	22.22 %
A2 level	66.67 %	11.11 %	22.22 %
B1 level	50%	28.57 %	21.43 %
B2 level	37.50 %	37.50 %	25 %
C1 level	63.64 %	36.36 %	0 %
Native speakers	0 %	57.14 %	42.86 %

Table 26. Mean percentage of production as a function of irregular no-change novel words divided by the level.

Language level	*skride*	*gleed*	other
A1 level	75 %	12.5 %	12.5 %
A2 level	35.71	0 %	64.29 %
B1 level	75 %	8.33 %	16.67 %
B2 level	16.67 %	33.33 %	50 %
C1 level	37.5 %	12.5 %	75 %
Native speakers	100 %	0 %	0 %

Table 27. Mean percentage of production as a function of selected no-change novel words divided by the level.

The common denominator of those five no-change past forms is their final alveolar phoneme: /t/ in the case of *drit* and /d/ in the case of *nold, gude, skride*, and *gleed.* The reason for the no-change past form may be (i) the analogy with other existing no-change forms such as *cut*, *put*, or *hit*, or (ii) the participants' unintentional feeling that the novel word already ends with an alveolar phoneme and no other past bound morpheme is needed since the bound morpheme blends with the final alveolar phoneme of the verb stem in their minds.

This type of irregular past-tense formation has also been previously mentioned in literature. Stemberger (1981) has shown that inflectional languages commonly avoid attaching "another" affix to a stem that already appears to have such an affix. This would explain why the unchanged past-tense forms originally ended in an alveolar phoneme. Bybee and Slobin (1982) conclude that it is not accidental that words whose past-tense form remains unchanged end their stem in either *t* or *d* (e.g., *hit* or *shed*) since the speakers may simply use a schema […t/d] $_{\text{verb/past}}$ in which the form of the stem and its past form would be identical. Berko (1958) has also described this phenomenon in her study, noting that the tested children were more reluctant to add a regular past-tense suffix to a novel verb that ended in an alveolar. Berko interprets this finding as an incomplete acquisition of the /ɪd/ allomorph. Bybee and Slobin (1982) counter this interpretation with the claim that such findings have more to do with the phonological shape of the novel form than with the late acquisition of the allomorph. Anisfeld and Gordon (1968) have shown that if the final phoneme of the stem shares some features with *t* or *d*, children are more likely to accept such novel forms as past-tense forms. Slobin (1971) pointed out that children also tend not to produce regular forms of verbs whose past-tense forms remain unchanged from the stem. Consequently, Kuczaj (1978) suggests that children are even more likely to accept unchanged past-forms of such verbs (and reject any such regularized forms, e.g., **hitted*). Both Berko (1958) and Bybee and Slobin (1982) then find a significant difference in this application between children and adults, who add the suffix more consistently. Bybee and Moder (1983) hypothesized that speakers tend to understand the verbs with the *-d* or *-t* stem ending as irregular in some respect and, therefore, supply past-tense forms identical to the stem. In their computer simulation of past-tense production, Rumelhart and McClelland's model (1986) also produced some strong no-change forms with verbs ending either in *-d* or *-t*. It is interesting to note that in the early stages of their simulation, the model was more prone to regularization (i.e., adding *-ed* to the stem); however, in the later stages, the no-change past-forms appeared as a part of the learning process. The authors also make an interesting note by pointing out that the words ending in *t/d* in the previous studies were all monosyllabic and it would be beneficial to see if the same phenomenon can be observed also in multisyllabic words such as *devote* or *decide*.

However, even though those five novel words show a certain tendency, it is also essential to mention the very low frequency of no-change past forms (irregular or not) in all the recorded responses.

4.4.3.2 Pronunciation of the bound inflectional morpheme

Another interesting feature noticed during the recording process is the full pronunciation of the final bound morpheme *-ed*, i.e., /ɪd/, in places where the phonological surrounding asks for a regular /t/ or /d/ pronunciation (after any phoneme except for the alveolar ones[40]) – for instance, the full pronunciation of *daped* as /deɪpɪd/ instead of /deɪpt/. The mean

[40] Instances of the full pronunciation after alveolars were not investigated as full pronunciation and counted under regular pronunciation since the full pronunciation of *-ed* after alveolars is considered phonologically correct.

percentage of its production was, therefore, calculated alongside the percentage of a regular pronunciation to see the ratio of the regular vs. full pronunciation. The production was also divided by the language level to see if this phenomenon possibly evolves over the process of language acquisition:

Language level	Regular pronunciation	Full pronunciation
A1 level	55.9 %	44.1 %
A2 level	83.65 %	16.35 %
B1 level	90.77 %	9.23 %
B2 level	94.74 %	5.26 %
C1 level	97.5 %	2.5 %
Native speakers	100 %	0 %

Table 28. Mean percentage of production of regular vs. full pronunciation of the morpheme -ed as a function of the language level.

The table above shows that the mean percentage of the full pronunciation in places where regular pronunciation should be is the highest (and also the most prominent) in the proficiently lowest A1 level and its distribution steadily declines with higher proficiency. Not surprisingly, native speakers did not produce any instance of full pronunciation. Similarly, the regular (i.e., correct, not full) pronunciation rises with higher proficiency, reaching almost 100%-production at the C1 level. Yet, it is interesting to notice that even the highly proficient C1 participants produced some past forms with the full pronunciation (while the native speakers did not). Another question thus arises whether its production could be phonologically conditioned for L2 learners, i.e., whether the full pronunciation could be triggered by the final phoneme of the preceding stem.

After analysing the phonological area preceding the past bound morpheme, i.e., the quality of the final phoneme of the stem, six types of word-final phonemes were identified:

1. plosives (i.e., /p/, /k/, /b/, and /g/ - excluding /t/ and /d/);
2. nasals (i.e., /m/, /n/, and /ŋ/);
3. liquids (only /r/, /w/ did not appear);
4. fricatives (i.e., /s/, /z/, /f/, /ʒ/ and /ʃ/);
5. affricates (i.e., /dʒ/ and /tʃ/);
6. and also collectively sibilants (/s/, /z/, /ʒ/ and /ʃ/, /dʒ/ and /tʃ/)

The distribution among language levels can be seen in the table below:

Level	Plosive	Nasal	Liquid	Fricative	Affricate	Sibilant
A1 level	40 %	12 %	4 %	34 %	10 %	41 %
A2 level	40 %	9 %	1 %	41 %	9 %	49 %
B1 level	40 %	6 %	6 %	40 %	9 %	45 %
B2 level	44 %	3 %	0 %	41 %	13 %	53 %
C1 level	38 %	0 %	0 %	50 %	0 %	44 %

Table 29. Mean percentage of production of a "full" bound morpheme as a function of the types of word-final morphemes divided by the language level.

The table above suggests that a big role is consistently played by stem-final plosives and sibilants (since fricatives can be almost wholly submerged under sibilants; the distribution of the non-sibilant /f/ with full pronunciation was the lowest of all fricatives). The high proportion of full pronunciation with stem-final sibilants could be explained by unintentional priming by a sentence preceding the sentence in the past: For instance, in *The baby likes to dize. Look, there he is dizing. Every day he dizes. So yesterday he…*, the full pronunciation of *-es* in *dizes* /daɪzɪz/ (which is phonologically correct after a sibilant fricative) could have primed the subsequent pronunciation of *dized* as /daɪzɪd/. The participants might have simply re-applied the full pronunciation to a wrong phonological position. The other phonological surrounding (i.e., plosives, nasals, and liquid /r/) also led to full pronunciations, though more rarely. These instances are very interesting in that they cannot be easily explained, referring to repetition; one of the possible explanations is that they are caused by hypercorrection. The L2 speaker may generally believe through a misunderstanding of the rules that the full form is preferable or formal (this could also be linked to Selinker's idea of the transfer of training (1972)). Another possible explanation is that the particular consonantal sequence may be hard to pronounce for the L2 speaker (since it is phonotactically illegal in their L2), and the learners may resort to the full form to make the pronunciation easier.

4.4.3.3 Inserted *s/t/d* phonemes

In the process of transcribing the produced past forms, an interesting and occurring phenomenon has arisen – the insertion of *-s-, -t-,* or *-d-* between the stem and the bound morpheme of past tense, e.g., *glips**ed* or *pank**t**ed*. Its frequency was very low (as seen in the table below); however, it was nonetheless interesting to look at its development over the process of L2 acquisition due to its recurring presence:

Language level	***-s-***	***-t-/-d-***
A1 level	2.76 %	1.57 %
A2 level	3.35 %	0.21 %
B1 level	0.7 %	0.17 %
B2 level	0.49 %	0.49 %
C1 level	0.31 %	0 %

Table 30. Mean percentage of production of inserted s/t/d morphemes as a function of the insertion type, divided by the language level.

The table above shows that even though the frequency of occurrence is rather low, there seems to be a progressive development from a higher frequency at the lowest levels (notice, for instance, the mean percentage of the inserted /s/ phoneme at the A1 and A2 level) to a minimal frequency at the C1 level. However, again, also the C1-level participants seem to be inserting the /s/ phoneme in-between the stem and the bound morpheme, and it is thus interesting to look at a possible stem-final phonetical conditioning in L2 learners (since the native speakers produced none of those insertion forms).

As for the full pronunciation phenomenon, four stem-final phoneme types were identified:

1. plosives;
2. nasals;
3. liquid /r/;
4. fricatives.

The tables below show the stem-final phoneme types and their mean distribution at the language levels. The first table shows the mean percentage of inserted *s*-phonemes, while the second table describes the distribution of *t/d*-phonemes:

Level	Plosive	Nasal	Liquid	Fricative
A1 level	71.42 %	7.14 %	14.29 %	7.14 %
A2 level	66.67 %	20%	6.67 %	6.67 %
B1 level	50 %	50 %	0 %	0 %
B2 level	66.67 %	0 %	0 %	33.33 %
C1 level	50 %	50 %	0 %	0 %

Table 31. Mean percentage of production of inserted s-phonemes as a function of the stem-final phoneme, divided by the language level.

Level	Plosive	Nasal	Liquid	Fricative
A1 level	37.5 %	37.5 %	12.5 %	12.5
A2 level	100 %	0 %	0 %	0 %
B1 level	100 %	0 %	0 %	0 %
B2 level	66.67 %	33.33 %	0 %	0 %

Table 32. Mean percentage of production of inserted t/d-phonemes as a function of the stem-final phoneme, divided by the language level.

Both tables show a rather prominent frequency of stem-final plosives and nasals that might trigger this specific phoneme insertion. The insertion of the *-s-* phoneme after these phonemes could be potentially explained by yet another unintentional priming (e.g., *glipsed* after *Every day he glips*) when the participant might have incorrectly use the inflected form *glips* as the stem. This insertion might also be implicitly conditioned phonologically by a number of similar phonemes attached one after another. In such cases, the speaker may use such insertion forms to differentiate the phonemes, a phenomenon known as dissimilation.

Such instances have also been previously described in classic literature as double past markers (in the case of *-t-* insertions) and as using affixes for both present and past (in the case of *-s-* insertions). In their computer simulation, Rumelhart and McClelland (1986), too, found the double past marking on seven responses to the input word. Even though they attribute such forms to errors that children and adults occasionally make, they also notice that the doubling occurred with those verbs whose stem ended in /p/ or /k/ and whose correct past-tense form should be created by the addition of a /t/ (e.g., /dript'd/). Interestingly enough, Berko (1958) also notes that several of the tested children retained the 3rd-person-singular /s/ or /z/ phoneme in addition to the past-tense suffix, as in /spowzd/.

4.4.3.4 Lexicalization of novel words

Similarly to the full pronunciation phenomenon, the recording process and the subsequent transcription of produced past forms discovered an occurring lexicalization of three types of novel words, specifically:

1. *bredge* – lexicalized into *bridged* in the past-form production
2. *drice* - lexicalized into *dressed* in the past-form production
3. *nace* - lexicalized into *niced* in the past-form production

The lexicalization of novel words has been previously reported by psycholinguistic studies (see, for instance, Sasisekaran et al., 2010)). However, the frequent occurrence of those three specific novel words and their past-form lexicalization appears rather peculiar. The table below was therefore used to see whether the lexicalization was a feature of L2 production only or whether the native speakers lexicalized as well and also subsequently whether the frequency of lexicalization evolved throughout the process of L2 acquisition. The table below, therefore, describes the ratio between the regular, non-lexicalized past-form and the lexicalized past-form, e.g., between *bredged* and *bridged*, divided by the language level.

Level	*bredged*	*bridged*	*driced*	*dressed*	*naced*	*niced*
A1 level	68.75 %	31.25 %	93.33 %	6.67 %	66.67 %	33.33 %
A2 level	69.23 %	30.77 %	91.67 %	8.33 %	69.23 %	30.77 %
B1 level	50 %	50 %	100 %	0 %	70.59 %	29.41 %
B2 level	72.22 %	27.78 %	92.86 %	7.14 %	80 %	20 %
C1 level	70 %	30 %	68.75 %	31.25 %	89.47 %	10.53 %
Native speakers	62.5 %	37.5 %	0 %	0 %	0 %	0 %

Table 33. The ratio between the regular (non-lexicalized) past-forms and the lexicalized past-forms, divided by the language level.

The table shows that in all cases, the regular past-form production prevails over the lexicalized past forms; however, its occurrence is by no means negligible. In all language groups, *bridged* seems to be the most prominent lexicalized past form of the three, followed closely by *niced* and *dressed* (in a reversed order for the C1 level). The native speakers only lexicalized *bredge.* Interestingly enough, the native speakers seem less prone to lexicalize the novel words than the L2 learners.

4.4.3.5 Item analysis

As a final step of the qualitative analysis, it was interesting to look at the differences in regular production of each novel word among language groups and between L2 learners and native speakers. Mean production of regular inflections of individual novel words was, therefore, calculated for the L2 learners to be compared with the production of the native speakers. The table below lists the novel words in order from the ones with the highest mean proportion of regular inflections (1), i.e., those that were in most cases produced as regular past forms, to the ones with the lowest mean proportion of regular inflections (32). This sorting allows for a row-to-row comparison between L2 learners and native speakers.

No.	L2 learners	Native speakers		No.	L2 learners	Native speakers
1.	*rask*	*pank*		17.	*plim*	*nold*
2.	*gezz*	*rask*		18.	*gare*	*shilk*
3.	*wiss*	*dape*		19.	*stin*	*tesh*
4.	*flidge*	*dize*		20.	*nung*	*teep*
5.	*stip*	*drice*		21.	*nold*	*blafe*
6.	*chake*	*gare*		22.	*bize*	*bredge*
7.	*tesh*	*nung*		23.	*stire*	*flidge*
8.	*glip*	*bize*		24.	*spling*	*plim*
9.	*pank*	*blig*		25.	*teep*	*stin*
10.	*dape*	*gezz*		26.	*drice*	*trisk*
11.	*blig*	*glip*		27.	*gude*	*gude*
12.	*dize*	*preak*		28.	*gleed*	*stip*
13.	*shilk*	*stire*		29.	*nace*	*drit*
14.	*trisk*	*wiss*		30.	*bredge*	*gleed*
15.	*preak*	*chake*		31.	*drit*	*skride*
16.	*blafe*	*nace*		32.	*skride*	*spling*

Table 34. vel words ranked according to their mean proportion of regular inflection with the focus on their different ranking between L2 learners and native speakers.

It can be deduced from the table above that some novel words received a similar ranking in terms of mean proportion of regular inflections for both L2 learners and native speakers (see for instance *rask*, *blig,* or *glip* for the higher proportion, and *skride, drit,* or *gleed* for the lower proportion). In other cases, the order of the novel words seems to differ quite prominently (see, for instance, *stip, tesh*, or *nung*).

4.4.3.6 Summary of the findings

The qualitative analysis has focused on some interesting phenomena noticed during the process of recording and transcribing. First, the analysis of past forms showed that at all language levels, the production of regular forms highly surpasses the production of irregulars, and that the production of regular forms declines with proficiency, while the production of irregular forms rises with higher proficiency.

The investigation of the types of irregular past forms showed that the internal change accompanied by the addition of a past bound morpheme *-ed* seems to be a dominant production pattern at the lower levels of L2 learners, and its use slowly declines with higher proficiency. At the lowest levels, this production is closely followed by irregular no-change forms whose production also shows a steady decline with higher proficiency. The internal change, on the contrary, is in the minority at the lowest levels and gathers noticeable strength with higher proficiency. The production of participants at the B2 and C1 levels seems to be closest to the native speakers.

The analysis of no-change past forms then showed that the recurrent no-change novel words have a common denominator, a final alveolar phoneme, and the no-change past forms may thus be the result of analogy or implicit processing of the final phoneme as the past bound morpheme.

The analysis of full pronunciation of *-ed* in places where the phonological surrounding asks for a regular pronunciation showed that the mean percentage of the full pronunciation is the highest and most prominent in the proficiently lowest A1 level, and its distribution steadily declines with higher proficiency. Nonetheless, its production seems to be phonologically conditioned since a big role is consistently played by stem-final plosives and sibilants. The high proportion of its occurrence after sibilants could be explained by unintended priming, while the other phonological surroundings are probably caused by hypercorrection. The L2 speaker may generally believe through a misunderstanding of the rules that the full form is preferable or formal. Another possible explanation is that the particular consonantal sequence may be hard to pronounce for the L2 speaker (since it is phonotactically illegal in their L2), and the learners may resort to the full form to make the pronunciation easier.

The investigation of the insertion of /s/, /t/, or /d/ between the stem and the bound morpheme of past tense showed a progressive development from a higher frequency at the lowest levels to a minimal frequency at the C1 level. It also hinted at a possible stem-final phonological conditioning in L2 learners, with a rather prominent frequency of stem-final plosives and nasals. The insertion of the *-s-* phoneme after these stem-final phonemes could be potentially explained by yet another unintentional priming (e.g., *glipsed* after *Everyday he glips*) when the participant might have incorrectly use the inflected form *glips* as the stem. This insertion might also be implicitly conditioned phonologically by a number of similar phonemes attached one after another. In such cases, the speaker may use such insertion forms to differentiate the phonemes.

Further analysis also showed an occurring lexicalization of *bredge, drice*, and *nace* in irregular past-form production. In all language groups, *bridged* seems to be the most prominent lexicalized past form of the three, followed closely by *niced* and *dressed*. The overarching item analysis showed some differences in production between L2 learners and native speakers. While the production of some novel words in past forms was rather similar (e.g., *skride* or *rask*), other novel words showed prominent differences in the mean production of regular inflections.

The qualitative analysis has then shown some specific pronunciation issues that need to be addressed didactically – e.g., the importance of proper pronunciation instruction of the *-ed* suffix, especially with complex consonant clusters and those that are phonotactically illegal in Czech. Pronunciation error correction should be consistent not only at the lowest levels but also with more proficient students to minimalize the incorrect pronunciation and ensure that such incorrect pronunciation does not stigmatize the L2 speakers in comparison to native speakers.

4.5 Discussion

The results of this experiment extend previous research (e.g., Albright & Hayes (2003), Ambridge (2010), and Blything et al. (2018)) to second-language learners of English and bring new insights into the acquisition of L2 morphology.

4.5.1 Summary of the findings

The analysis of the produced forms has shown that for the native speakers, the novel verb's similarity to existing regular verbs played a significant role in the production of past forms, in line with the single-route model. However, L2 learners have shown lesser dependency on verbal similarity to regulars when deciding on its past form. Further post-hoc tests have shown that the way in which L2 learners acquire the English past tense is characterized by a progressive development from the mastery of mechanistic rules at the A1, A2, and B1 levels to the refinement of their application by spotting analogical patterns of existing verbs at the B2 and C1 levels, similarly to native speakers. The second-language speakers thus show changes in the development of inflectional morphology that come closer to adult native speakers with the higher proficiency of B2 and C1 levels, while lower levels seem to adopt an alternative strategy.

Further analysis of the reaction times has also shown that for the A1-level learners, alongside the A2-level learners (although with a weaker effect), Czech implicitly functions as a facilitator in their performance, with participants being faster with novel words that are phonotactically legal in both Czech and English. L2 speakers with higher proficiency and native speakers, instead, did not show any effect of Czech phonotactic probabilities. The qualitative analysis has also shown some specific pronunciation issues, such as the use of full pronunciation of bound morphemes where a single phoneme is expected that need to be addressed didactically. It has also revealed the insertion of additional phonemes in the bound morpheme, possibly due to priming.

4.5.2 Discussion and conclusion

Recent production tasks using novel words (e.g., Albright & Hayes, 2003; Ambridge, 2010 or Blything et al., 2018) have provided some evidence that the production of the English past tense shows signs of analogical processes associated with a single-route model. So far, the role of the similarity to existing regular verbs on the production of novel verbs has been studied both in native English-speaking children (Ambridge, 2010 and Blything et al., 2018) and adults (Albright & Hayes, 2003; Ambridge, 2010), and also in L2 learners of English (Beck, 1997; Ellis & Schmidt, 1997; Murphy, 2000; Agathopoulou, 2009; Cuskley et al., 2015). Our study further investigates whether evidence for the analogical morphological productivity extends even to L2 learners of English with Czech as L1 with varying proficiency levels and whether the process behind the production differs among levels, starting with a proficiency sample that has just started acquiring English morphology and moving to the language sample closest to English native speakers.

Our results suggest that the lowest proficiency groups (A1, A2, and B1) generate regular past forms by the default rule of adding *-ed* to the novel verb's stem, seemingly irrespective of how similar the novel verb was to a real English regular verbs. In contrast, the more proficient groups (B2 and C1) and native speakers of English already seem to have developed a morphological system in which the production of regular past forms depends on how phonologically similar the novel verb is to existing verbs based on analogy. These findings extend previous findings that spoke in support of analogical processes associated with morphological productivity (e.g., Albright & Hayes, 2003, Ambridge, 2010, or Blything et al., 2018) to more proficient levels of second-language learners of English. Blything et al. (2018: 14) use their findings to show that "overregularization errors made by children […] need not be attributed to a rule-based mechanism […] and, indeed, are better explained in terms of analogy across stored exemplars." Similarly, our results suggest that this claim could be extended to the L2 learners at the B2 and C1 levels.

Such a cut-off point nicely corresponds to the proficiency differences found in Experiment 1, where the A1 and A2 levels performed differently from the B1, B2, and C1 levels. However, in Experiment 1, all language levels preserved the general pattern of reaction to individual conditions. In this experiment, the similarity to existing (ir)regulars turned out to be a significant predictor only for the more proficient learners. Also, in Experiment 1, both native speakers (tested by Cilibrasi, 2016) and L2 learners preferred morpheme stripping, L2 learners then in combination with analogical processing based on phonotactic probabilities of presented stimuli. This finding suggests that in perception, a dual (rule-based decomposition and analogical processing) route mechanism (or a mix of decompositional and atomist theories) is used from low proficiency onward. On the other hand, Experiment 2 shows that in production, there is a transition from the dual-route (at A1 to B1 levels) to single-route model (at B2 and C1 levels and native speakers). Since the first experiment is not focused on an "active use of the language" (contrary to the production experiment), its findings may be more strongly influenced by the morphological processes that are active in the participants' L1 (e.g., by the heightened morphological awareness of speakers of a morphologically rich L1). The results of the production experiment, on the other hand, may be the consequence of the teaching of English morphology in Czech schools at the lower levels (where especially *-s* and *-ed* endings are continuously drilled). In comparison, learners with higher proficiency are exposed to more English language in general and seem to start understanding the patterns in the language and apply them analogically to new units. The lack of the effect of the similarity to existing regulars might also be influenced by the lack of language knowledge at the lower levels (when rule application might be more automatic and easier than applying analogical patterns) and insufficient language input for implicit learning.

Following the so-called entropy theory (e.g., Lõo et al., 2018), the production of English past-tense forms by L2 learners can be influenced by the number of past-tense morphemes both in English and Czech, their frequencies, and phonological similarities. English regular and irregular past tense forms have different probability distributions in English L2 speakers due to the difference in their informational content. The fact that our lower-proficiency participants tended to primarily use the regular past-tense rule might be explained by their primary familiarity with the regular past-tense morphemes (especially at the A1 level). They might be more prone to produce only regular past-tense forms since irregular past-tense forms are less familiar and, therefore, have less probable active distribution. As L2 learners are exposed to more significant language input, the entropy values for both regular and irregular past-tense forms might change. This claim nicely fits our data, as the higher-proficiency participants showed lesser dependence on the rule application and started using analogy instead.

In their study, Blything et al. (2018) make a thick dividing line between the effects of the single- and dual-route model, trying to distinguish the former from the latter. However, given the results of our research that points to a gradual progression from rule application to the use of analogy, and in line with previous research based on seemingly incompatible models (e.g., Cilibrasi et al., 2019), this discussion part attempts to propose a less dividing conclusion (in line with our previous argument in Chapter 3). Redundant models that describe the processing of morphologically-inflected forms particularly well, such as the one proposed by Schreuder et al. (1999), hold that such processing generally involves two types of analyses that seem to work simultaneously: single- (rote-based parsing) and dual-route mechanisms (rule-based parsing). Those analyses operate in parallel, but each system has varying importance connected to how frequently a given item is generally used. A mechanism of this kind may be at play in second-language learners of English. Albright and Hayes (2003: 120) themselves argue for a new (third) "model of morphology that makes use of multiple, stochastic rules," and our results suggest that these rules may equally apply to L2 learners, with the combined contribution of simple rules and generalizations based on analogy. These two mechanisms may be present at all levels but have varying importance depending on proficiency.

The idea of a redundant system is not new in the field. In an influential publication, Alegre and Gordon (1999) question the distinction of irregular verbal forms being stored in our memory and of regularly inflected forms never being represented in such a way, stating that generalization based on rules by no means exclude rote-learned representations. In their article, the authors refer back to Sandra (1994), who already pointed out that rule-based production of inflectional morphology entails great processing demands, and the combination of two processes makes the production of inflected forms more convenient (Alegre & Gordon, 1999). While "the strong version of the dual-route model" proposed by Prasada and Pinker (1993: 9) accepts no whole-word representation for the regularly-inflected verbs in memory, some storage (though not necessary) is "made possible" in the weaker form of the model they consequently proposed. The authors state that "prior storage of regulars is possible, and thereby might offer mild analogical assistance to the generalization of the regular inflection to similar forms" (Prasada & Pinker, 1993: 9). Pinker and Ullman (2002: 458) further stress out that the dual-route theory "does not posit that regular forms are never stored, only that they do not have to be," proposing that both rules and analogy might be involved in how regular past forms are generated. In this article, the authors give the example of verbs with both regular and irregular past-tense forms (for instance, *dreamed – dreamt*), with which the storage of the regular past-tense form would be needed to avoid blocking by the irregular form. This proposal goes hand in hand with the previous findings of Alegre and Gordon (1999) that a limited number of frequently-used items would be stored directly in memory. However, this version of the model raises a number of questions in terms of storage and verb activation since some of its features are vague. Blything et al. (2018: 3; 14) point out that the weaker version of the dual-route model is "empirically indistinguishable from the single-route model," "with a default rule bolted on (and no clear specification of when this rule should be used instead of analogy)."

Pinker and Ullman (2002: 462) also note that the scientific debate has been too focused on finding the one right model that would explain the production of morphologically inflected forms, to the point that it might "have prevented both sides from acknowledging that features of one model may correspond to constructs of the other, described at a different level of analysis." They propose letting the data show us "which processes are most naturally explained by which kinds of mechanisms" (Pinker & Ullman, 2002: 462) instead of attempting to subsume all the phenomena under one of the two mechanisms. The current study perfectly fits this approach and shows that both models may have a role in explaining second language acquisition of morphology, with one model describing the early stages and another model describing the later stages of English past-tense acquisition.

In conclusion, the second experiment has enabled us to look more closely into the processes underlying morphological productivity in second-language learners of English and also into its potential development with higher proficiency. The most important contribution of this experiment lies in showing that the L2 acquisition of past-tense morphology is characterized by a gradual progression from the application of default rules to the application of analogical patterns.

5. EXPERIMENT 3: THE LEXICAL DECISION EXPERIMENT[41]

5.1 Introduction

The reaction-times analysis in the previous chapter has shown that the A1-level learners, alongside the A2-level learners (though with a weaker effect), reacted in a significantly quicker way to the words phonotactically legal in Czech. This indicated that Czech functioned as a facilitator in their performance. For the B1, B2, and C1 levels, and also for the native speakers, Czech had no effect on production since participants reacted similarly to novel words of both sets. This suggested that for the more proficient levels and native speakers, there were no transfer effects at play, and Czech functioned neither as a facilitator nor as an inhibitor. These findings hint at the fact that for the lowest L2 levels, the mother tongue still plays some role in the usage of their L2 and can facilitate the learning process. With the more proficient levels, the L2 starts to be a unit of its own, seemingly independent of the L1 influence.

This final experiment is designed to supplement the previous production experiment in an attempt to understand any other possible transfer effects behind the analysed novel words and subsequently behind L2 production and perception. As such, it may bring more insight into how (or if) L1 affects the acquisition of L2 since a crucial debate in L2 learning and bilingualism research concerns the access to the lexicon. The experiment employs the use of real English words and phonotactically-legal English novel words (half of which is phonotactically legal and a half of which is phonotactically illegal in Czech) to help explore the process of L2 development in relation to L1 and potential transfer effects.

5.2 Theoretical background and hypotheses

5.2.1 Theoretical background

Second-language learning generally aims to enable learners to use the L2 without the mediation of their L1; however, suppressing one's mother tongue's interference is one of the most difficult challenges for an L2 learner. L1 transfer seems to affect almost all aspects of L2

[41] Parts of this chapter will be hopefully published (with some changes) in *English Teaching & Learning*. The paper was under review at the time of the thesis submission.

learning, including phonology, syntax, and spelling (Darcy et al., 2018). Previous studies have found evidence for L1 transfer onto L2 across many language areas, and the initial reliance on the L1 does not seem to develop into a completely independent use of L2 even with higher L2 proficiency. Recent findings show an effect of L1 on L2 processing, even in proficient bilinguals (Kroll et al., 2002). The role of L1, therefore, does not seem to be easily separated from the processing in L2, and it is necessary to study its effect on the whole process of L2 acquisition. One interesting domain for investigating this transfer is the lexicon: How do L2 learners acquire representations for L2 words in connection to their L1 representations of word meaning? How are the L2 words stored and processed, and do the learners suppress their L1 in the process?

Research on transfer effects generally offers two alternatives of how the two languages interact in the mental lexicon. One of the alternatives is the word association model, which posits that "a direct association is established between words in the two languages" (Potter et al., 1984: 23), meaning that "L2 words are mediated via direct connection to their translation equivalents in L1" (Kroll et al., 2002: 137). The other alternative is the concept mediation model, which "proposes that the only connection between the two languages is via an underlying, amodal conceptual system" (Potter et al., 1984: 23), meaning that "L2 words are connected directly to their meanings without L1 mediation" (Kroll et al., 2002: 137–138). In their study, Potter et al. (1984) suggested that the concept mediation model was used even by less proficient learners. However, subsequent studies have shown that at the beginning of the learning process, L2 learners seem to depend more on the word association model (and therefore using the L1 mediation for L2 words), while with later stages of proficiency, the learners seem to gradually move onto the concept mediation model and side-line their L1 in favour of making a direct connection between the concepts and their L2 realizations (Kroll et al., 2002). L2 learners, therefore, seem to decrease their dependence on L1 with higher proficiency and access L2 words directly (Kroll et al., 2002). In connection to the relation between the two lexica and their storage, van Heuven et al. (1998) proposed a distinction between language-dependent lexica in which "bilinguals possess one integrated lexicon for both their languages," and language-independent lexica in which bilinguals operate with "two separate lexica, one for each language" (Van Heuven et al., 1998: 458).

Concerning the online processing of words from the two lexica, another concern of transfer studies lies in the dilemma of how the two lexica are accessed, and more specifically, on whether the access to the two languages is selective or non-selective. The problem revolves around the following questions: When perceiving or producing words in one language, do L2 learners activate words in both languages that they have a command of (as proposed by the language non-selective access view), or do only words of that specific language triggered by the stimuli get activated in isolation (as proposed by the language selective access view) (de Bot et al., 1995)? Dijkstra & Kroll (2005) claim that bilinguals generally believe they activate only the target language in language processing (for further evidence on language selectiveness, see, for instance, Caramazza & Brones, 1979). However, several research findings show that during language processing, L2 learners seem to active the lexica of both languages in non-selective language access (Van Heuven et al., 1998; Lemhöfer & Dijkstra, 2004; Brenders et al., 2011). To illustrate this simultaneous activation, we can refer, for instance, to Van Heuven et al.'s (1998) study, which showed that bilingual participants are affected by orthographic neighbours of both languages when recognizing target words of only one language, showing that both lexica must have been activated simultaneously.

To investigate the (in)dependence and (non)selectiveness of lexica in the language processing of L2 learners, we decided to study how L2 learners of varying language proficiency process novel English words that were either phonotactically legal or illegal in their L1. For these purposes, we employed an auditory lexical decision task to study the organization of their lexica. In its basic form, the task is used to "present auditory input to participants who have to decide whether what they hear is a word or nonword and respond by pressing two buttons, one for word and the other for nonword. They have to respond as quickly and accurately as possible" (Jiang, 2013: 86). While second-language research has made use of lexical decision tasks before, most researchers have employed a written word recognition methodology, and auditory lexical decision tasks have been rather overlooked (Jiang, 2013). Jared and Kroll (2001) used this method to study interference effects in French-English bilinguals, noting L1 interference for French words that showed similarity to English words and phrases. Marian et al. (2008) used the same method to study cross-language phonological overlap between English native speakers and Russian L2 learners of English. Their L2 participants reacted faster and with greater accuracy to English word stimuli that were phonological similar to Russian, as opposed to words with no Russian influence and English-specific phonemes. Phonological similarities between the two languages, therefore, seemed to facilitate auditory word recognition in L2 learners. These findings are in line with Jared and Kroll's study (2001), which also showed a facilitatory effect of L1.

Previous studies have also shown that proficiency is a decisive factor in the organization of the two lexica. Kroll (1993) reported that with increasing proficiency, there comes a transition from lexical to concept mediation (i.e., from the word association model to the concept mediation model). Kroll and Curley (1988) found evidence for concept mediation in German L2 learners only after two years of classroom learning. Talamas et al. (1999) found that words with a similar form in both languages caused more interference in less proficient L2 learners. This finding has been further supported by Jared and Kroll's study (2001) on English-French L2 learners. Kroll et al. (2002) have shown that their L2 English-French learners with higher proficiency were faster in naming words in L2 and translating between the languages. Harrington (2006) used a lexical decision task to measure L2 lexical proficiency, showing that both reaction times and accuracy decreased systematically with increasing proficiency and the frequency of the presented stimuli word. However, very little information is given on how exactly the native language influences the L2 acquisition in its full length, i.e., not only at the beginning of L2 acquisition but also in the later stages and even at the top of proficiency. This led us to investigate the effect of L1 on English novel words in L2 learners of English at different proficiency stages.

5.2.2 Hypotheses

Our starting hypotheses were built on the findings of the previous experiment. We expect the reaction time to real words to be quicker than to the novel words since the real words will sound more familiar to the participants and facilitate a prompter reaction. We also expect the set (i.e., the novel word's phonotactic il/legality in Czech) to be a decisive factor in the quickness of the participants' reaction or in their accuracy, with the L1 facilitating the task (based on the studies mentioned above), at least for some of the L2 levels (since the lower levels might suffer from more interference).[42] The participant's score in the placement test may also potentially affect the performance, promoting a quicker or more accurate response with higher proficiency.

[42] Here, we understand language transfer as positive and facilitating language learning, in contrast to interference, which is understood as a negative type of transfer that potentially hinders language acquisition (see, for instance, Gao, 2013).

5.3 Methodology

5.3.1 Ethics

The experiment was approved by the Charles University Ethics Committee since it met all the requirements on how the data were anonymized and how the personal data were handled.

5.3.2 Recruitment

The participants were recruited via an e-mail. Only second-language learners of English were recruited since recruiting native speakers for the decision between real words and novel words would not be effective for our purposes, and the findings of the previous experiment have also shown that for native speakers, there is no effect of Czech underlying their production speed.

The English second-language learners with L1 Czech were again recruited from the Faculty language school. The testing group comprised a selection of the participants that completed the previous production experiment. However, special attention was paid to the necessity of a long-time span between participating in Experiment 2 and this experiment to prevent the remembrance of the novel words. For this reason, the testing for the third experiment started more than six months after the previous testing for novel-word past-tense production.

Participants with a required placement-test percentage (see the table below) were addressed directly. A small fee was given to all participants (as this research project has been financially supported by a Specific Academic Research grant, "Jazyk a nástroje pro jeho zkoumání").

5.3.3 Participants

30 English second-language learners with Czech as their first language were recruited. None of the participants had been diagnosed with any language impairment or hearing deficits. The mean age was 36 years; 23 participants were female, 3 were male.

In contrast to the previous production study, in this experiment, the participants were carefully selected by their overall performance in language placement tests[43] to create a continuum scale of language proficiency, from a low language score (21%) to high proficiency (94%), as assessed by an English language assessment test. We aimed at a strong linear correlation between the participant's rank and their language proficiency (see the table below). With this design, we can observe how the potential transfer effects, and more generally, how the performance gradually changes with proficiency.

Rank	Total score	Rank	Total score	Rank	Total score
1	21%	**11**	48%	**21**	73%
2	23%	**12**	52%	**22**	75%
3	25%	**13**	54%	**23**	79%
4	26%	**14**	57%	**24**	81%
5	28%	**15**	59%	**25**	85%
6	30%	**16**	62%	**26**	87%
7	32%	**17**	64%	**27**	89%
8	35%	**18**	67%	**28**	91%
9	40%	**19**	69%	**29**	93%
10	43%	**20**	71%	**30**	94%

Table 35. The proficiency scale based on the correlation between rank and placement-test scores.

The scores were mainly kept 2-3% apart, with no more than 5% when there was no other participant to recruit for the study with a similarly suitable score. The correlation between the rank and the score amounted to 99.69% (a nearly perfect correlation). The distribution can be observed in the figure below:

[43] Having counted a mean of test scores across individual language-level placement tests, all levels have been intrinsically assigned the same weight, while some might argue that the more difficult placement tests should have higher weights. However, it is not an easy question to decide which weight they should have and how this could be incorporated into creating a proficiency scale. Therefore, assigning the same weight to the individual placement tests seemed as the most systematic solution. The more proficient B2- and C1-level tests nevertheless had 30 questions instead of 20 (as in the case of the A1, A2, and B1 level), which could satisfy the weighting argument.

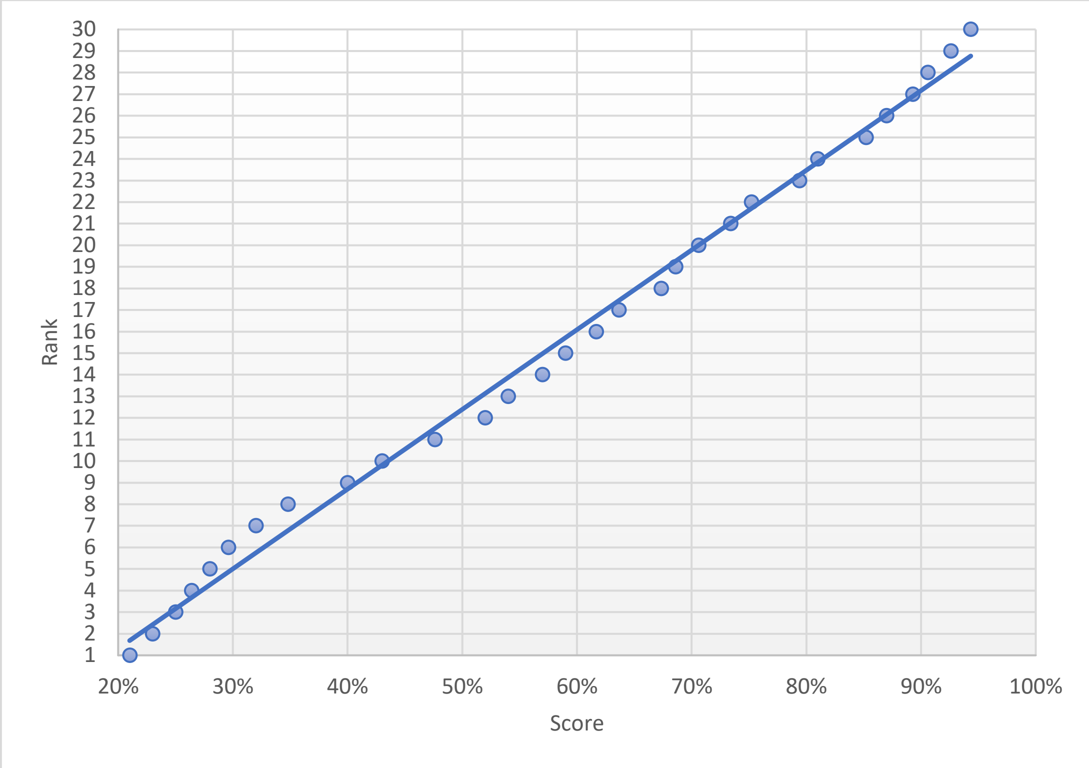

Figure 22. The rank-score correlation for the proficiency scale.

By distributing the language performance on a proficiency scale (see, for instance, Crisp & Lambon-Ralph, 2006), this experiment attempted to explore potential proficiency effects on the performance through the use of a continuous proficiency measure rather than dividing the proficiency into separate categories such as language levels (see also Cilibrasi et al., 2019). This method allowed us to keep all L2 learners in one sample and use the score as one of the performance predictors.

5.3.4 Consent

Similarly to the previous experiment, the testing took place in a quiet room located at the Faculty of Arts. The participants werethen asked to express their consent with the experiment, data anonymization, and research goals. They were also informed that all the data would be stored in a locked cabinet to which only the administrators would have access. Each participant was then given a random five-digit number[44] in order to anonymize their data.

[44] Generated at https://www.random.org/.

5.3.5 Stimuli

The 32 novel verbs were adopted from Experiment 2 (for the full description, see Chapter 4 above). To create a balanced dataset and provide a counterpart to the novel words, 32 real English words were added to the task. Those were imported from the British National Corpus (BNC) accessed through korpus.cz and its search engine KonText (Machálek, 2020),[45] using a CQL query [tag=VVB][46] with lemmas, filtering 1,416,263 hits, which included 5,440 different items. Those had to be checked manually so that the verbal lemmas did not function as converted nouns or adjectives. The items were then filtered by lemma frequencies. Since the novel words were all entirely monosyllabic, the real words had to be monosyllabic as well so that the difference in syllables did not influence the quickness of reaction. The real words were also chosen so that 11 of them were of the highest frequency in the corpus, 10 of them of medium frequency, and 11 of them of the lowest frequency (to avoid favouring high-frequency real words in contrast to unknown novel words, which might have influenced the results). For the complete list of the stimuli, see Appendix 3.

5.3.5.1 Stimuli recording

The stimuli were recorded simultaneously with the stimuli for Experiment 2 in the sound booth of the Institute of Phonetics at the Faculty of Arts, Charles University, by a native speaker of English who had been previously instructed to record the frame sentences with proper intonation and a special focus on the novel word.

5.3.6 Procedure

Following the lexical decision paradigm, each participant was presented with those 32 English novel verbs and 32 real words. They were told to decide whether the word they heard in each trial was a real English word or not by pressing one of the assigned keys on the keyboard when the recording stopped playing.

[45] https://www.korpus.cz/kontext/query?corpname=bnc. Last accessed August 16, 2018.
[46] The *VVB* tag returns all finite lexical verbs, excluding all gerund forms, modal and auxiliary verbs (see http://www.natcorp.ox.ac.uk/docs/c5spec.htm).

The experiment was created in the software PsychoPy v3.0 (Peirce et al., 2019). Each participant was first presented with one practice trial with one stimulus to get accustomed to the task. The trial session was followed by a test session, in which the participants were presented with the lexical decision task. In total, the participants heard 64 words, half of them were real English words, and the other half were phonotactically-legal English novel words. All stimuli were presented randomly, and the order of stimuli presentation was different for each participant. Reaction times and pressed keys were recorded.

5.4 Results

5.4.1 Reaction-times analysis

The reaction-times analysis is designed to answer whether reaction times are different for novel words phonotactically legal and illegal in Czech and whether proficiency has any role in performance. Its findings would offer us a closer look into possible transfer effects at play during lexical access.

A total of 1,875 reaction times were recorded. The trials were first checked for outliers (for our definition of an outlier, see Chapter3). For the presented data, the first quartile (FQ) starts at 0.44 seconds, and the third quartile (TQ) ends at 0.85 seconds, thus creating an interquartile range (IQR) of 0.41 seconds. Using the definition of an outlier, the upper bound for the regular data equals 1.5*(TQ + IQR), i.e., 1.47 seconds. The lower bound then equals 1.5*(FQ – IQR), i.e., -0.17 seconds. All data points that do not fit into the range of -0.17 and 1.47 seconds were then considered outliers and deleted from the final data set. Since our primary interest was the speed of reaction with a correct response to a given stimuli word, the reaction times for incorrect answers were excluded from the dataset as well.[47] After the omission of the outliers and incorrect responses, the dataset had the following characteristics:

[47] More on this exclusion in the discussion section below.

	N	Mean RT (s)	Min. RT (s)	Max. RT (s)	Standard deviation (s)	Mean RT (real words; s)	Mean RT (novel words; s)	Mean RT (A set; s)	Mean RT (B set; s)
RTs	1,313	0.66	0.02	1.45	0.28	0.58	0.74	0.72	0.75

Table 2. Descriptive statistics of the dataset used for the reaction-times analysis, including the number of items, mean reaction times, minimum and maximum reaction times, standard deviation, and mean reaction times of the real/novel words and the A/B sets.

For each trial, reaction times were recorded (along with the ID number, the placement test score, the item type (real vs. novel word), and for the novel words, even the set (A – phonotactically legal in Czech – or B – phonotactically illegal in Czech).

Given the size of the analysed sample that consists of 30 participants, the analysis assumes normal distribution based on the central limit theorem (see, for instance, Kwak & Kim, 2017; or above in Chapter 3) and, therefore, employs parametric tests. To investigate whether the data showed significant effects of various independent variables on reaction times, the results were once again analysed with a linear mixed-effects model[48] in R (R Development Core Team, 2011) to investigate manipulated variables and random effects that may appear as a result of individual differences among participants since the model works with individual trials rather than average scores (Baayen et al., 2008). The dependent variable used was the reaction times to the presented stimuli; the independent variables included (a) the score, (b) the word type (a real or novel word), and (c) the set (A or B in novel words). All those variables were entered into the model at the same time to investigate their possible interaction. To account for possible multicollinearity of the independent variables, the variance inflation factor (VIF) was checked for each model. The values for the reaction-times analysis were as follows: type - 1; score - 1; set – 1. The model accounted for random effects by keeping their structure its simplest and most-frequently-used form of (1|part) + (1|item) with "part" standing for participants and "item" for individual stimuli words. As stated above, only reaction times with correct answers were used in the analysis.

In the initial analysis, we investigated the whole dataset, i.e., both real and novel words, to determine whether novel words are processed differently from real English words and whether language proficiency influences the performance in any way. We started the analysis with a two-way interaction of word type and score; the p-values were generated by the *lmerTest*

[48] The data meet all the necessary assumptions of the model.

package (see Kuznetsova et al., 2012). The overview of the linear model can be seen in Appendix 3. To get a simpler overview of the significant variables, an analysis of variance (ANOVA) was then run on the model. The analysis showed a significant effect for the word type, but not for the score, as seen in the table below:

	Df	Sum Sq	Mean Sq	F value	P-value
Type	1	0.453	0.453	8.429	**0.004 ****
Score	1	0.044	0.044	0.812	0.375
Type:Score	1	0.013	0.013	0.233	0.630

Table 36. The analysis of variance (ANOVA) table for the initial reaction-times analysis. The numbers in bold indicate a significant effect. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Those results suggest that the participants reacted differently to the two types of presented stimuli with no significant difference in relation to proficiency scores. Therefore, the difference in reaction speed to the two stimulus types was relatively constant for all participants and all language scores. A graphic representation of the finding is provided below:

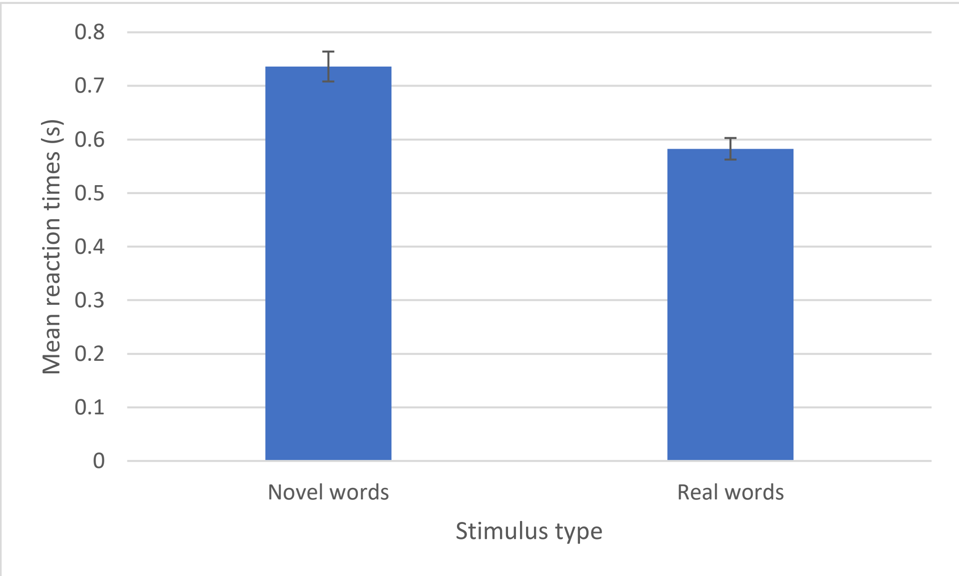

Figure 23. Mean reaction times for real and novel words.

The figure above shows that the participants took longer to discriminate novel words in comparison to the real words, which took significantly less time to discriminate.

In the follow-up analysis, we focused only on the investigation of the novel words to determine whether English novel words that are phonotactically legal in Czech are processed differently from English novel words that are phonotactically illegal in Czech and whether language proficiency influences the performance as well. In other words, we aimed to investigate whether Czech plays any role in L2 word recognition (and, if so, whether Czech facilitates or inhibits the reaction) and whether its role differs for different language proficiencies. Again, we started the analysis with a two-way interaction of score and set (A or B based on the novel word's phonotactic il/legality in Czech). The overview of the linear model can be seen in Appendix 3. To get a simpler overview of the significant variables, an analysis of variance (ANOVA) was then run on the model. The analysis showed that none of the variables is a significant predictor, neither is their interaction, as seen in the table below:

	Df	Sum Sq	Mean Sq	F value	P-value
Set	1	0.007	0.007	0.125	0.724
Score	1	0.014	0.014	0.237	0.630
Set:Score	1	0.058	0.058	0.972	0.325

Table 37. The analysis of variance (ANOVA) table for the advanced reaction-times analysis.

Those results suggest that the participants of all proficiencies reacted similarly to both sets' novel words and, therefore, made no difference between novel words phonotactically legal or illegal in Czech. In this respect, it would seem that Czech functions neither as a facilitator, neither as an inhibitor, not affecting the speed of the reaction. However, when mean reaction times were counted for the A and B subset for each participant and an independent samples t-test was run between those two subsets (mean A and B reaction times) without random effects, its results turned out marginally significant, two-tailed: $t(29) = -1.82$, $p = .079$ (significant when one-tailed: $t(29) = -1.82$, $p = .04$). These findings suggest that some transfer effects might be at play, but their strength might be simply relatively weak. Plotting was needed to understand the nature of such an effect:

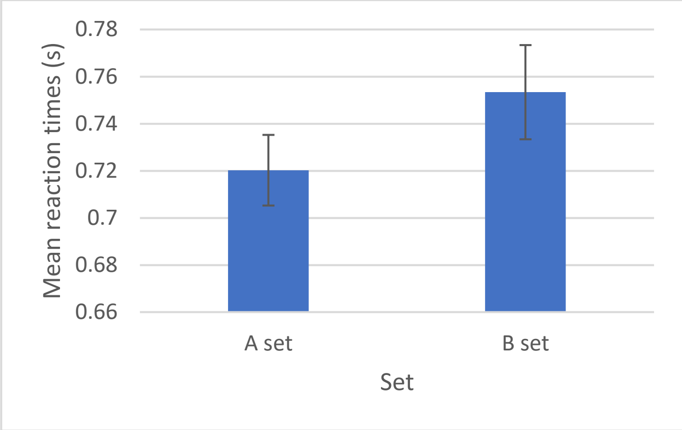

Figure 24. Mean reaction times to the A and B set.

The figure shows that the participants took more time to discriminate the novel words of the B set that are phonotactically illegal in Czech. However, as previously stated, the results of the independent samples t-test turned out only marginally significant, and its interpretation must be made with extreme caution. It is nevertheless possible that Czech has a weak effect on L2 perception.

5.4.2 Accuracy analysis

The accuracy analysis attempts to answer whether the correctness of an answer to a given stimulus was influenced by the type of the word (real vs. novel), its set (A or B), or the proficiency score. The findings would offer us further information on possible transfer effects at play during lexical access.

A total of 1,875 answers were recorded. The mean proportion of correct answers for each word type (real or novel) and in novel words also for each set (A or B) is shown in the table below.

	N	Mean accuracy	Mean accuracy (real words)	Mean accuracy (novel words)	Mean accuracy (A set)	Mean accuracy (B set)
accuracy	1,875	69%	68%	70%	75%	66%

Table 38. Descriptive statistics of the dataset used for the accuracy analysis.

For each trial, key presses were recorded (along with the ID number, the placement test score, the item type (real vs. novel word), and for the novel words, even the set (A – phonotactically legal in Czech – or B – phonotactically illegal in Czech).

To investigate whether the data showed significant effects of various independent variables on the chosen answer, the results were analysed with a generalized mixed-effects model.[49] The dependent variable used was the accuracy (i.e., whether the answer was correct or not) to presented stimuli; the independent variables included (a) the score, (b) the word type (real or novel words), and (c) the word set (A or B in novel words). To account for possible multicollinearity of the independent variables, the variance inflation factor (VIF) was checked for each model. The values for the reaction-times analysis were as follows: type - 1; score - 1; set – 1. The model used participants and items as random factors (the most straightforward and frequently-used code for random effects as described above in the reaction-times analysis).

The initial analysis focused again on the whole dataset to see whether the accuracy of performance differed between real and novel words and whether language proficiency influenced the level of accuracy or not. As in the previous analysis, we started with a two-way interaction of word type and score. The overview of the generalized linear model can be seen in Appendix 3. To get a simpler overview of the significant variables, an analysis of variance (ANOVA) was then run on the model. The analysis showed a significant p-value for the interaction between the word type and language score, but not for the type and score individually, as seen in the table below:

	Df	Sum Sq	Mean Sq	F value	P-value
Type	1	0.02	0.02	0.02	0.888
Score	1	2.60	2.60	2.60	0.118
Type:Score	1	7.97	7.97	7.97	**0.009** **

Table 39. The analysis of variance (ANOVA) table for the initial accuracy analysis. The numbers in bold indicate a significant effect. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Those results suggest that the participants were similarly accurate when presented with the real and novel words, without any significant difference between the two types of stimuli. The results of the ANOVA also did not uncover any significant differences in proficiency scores, suggesting that participants of various language scores performed with a similar level of

[49] The data meet all the necessary assumptions of the model.

accuracy. The significant interaction between the word type and language score suggests that the accuracy to the two stimulus types varied for different language scores. Post-hoc group t-tests were, therefore, run to understand the direction of this interaction. The post-hocs showed that the score significantly influenced accuracy only in real-word stimuli (with p-value .048) but not in novel words (p = .703). To understand this interaction better, graphical visualization is provided below:

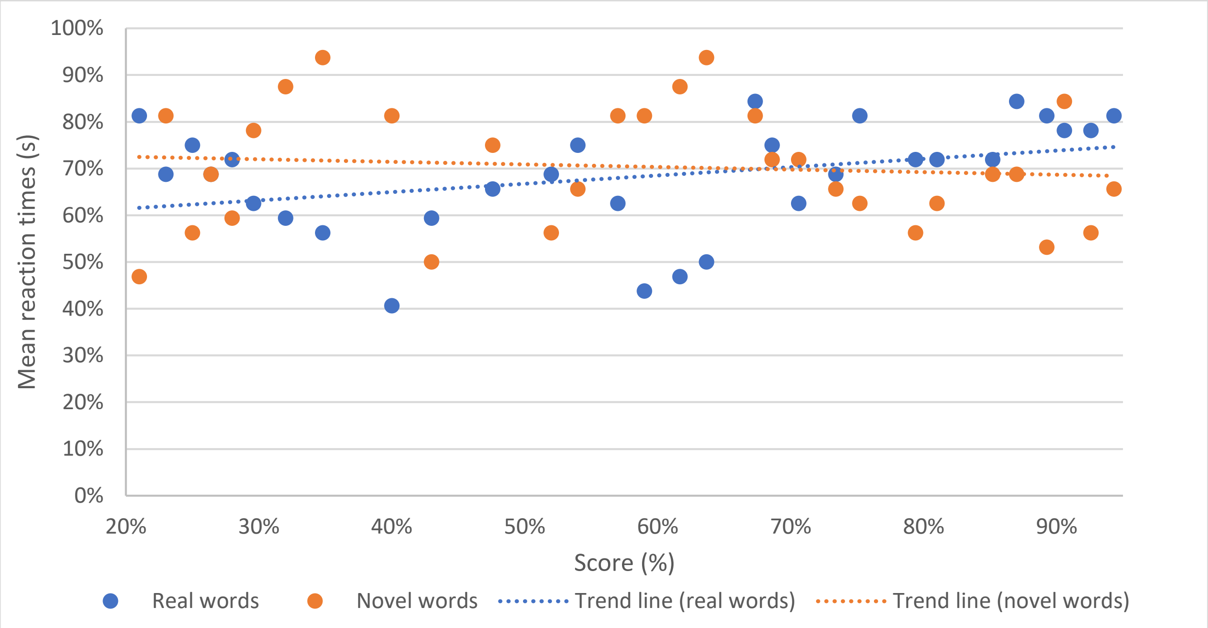

Figure 25. Mean accuracy for real and novel words on the proficiency scale.

The plot shows that the accuracy trendline for novel words remains relatively constant across the span of the score scale (the slope is virtually flat). The participants were, therefore, reacting to the presented novel words with a relatively similar level of accuracy, were they of a low or a more proficient language score. On the contrary, accuracy to the presented real-word stimuli shows a progressive development with raising proficiency. More proficient participants reacted more accurately to the presented real words than lower-proficiency participants. Those findings can be explained by greater familiarity with the language and a potentially larger lexicon in advanced speakers. It is reasonable to assume that low-frequency real words are recognized with greater accuracy in more proficient L2 learners, explaining the pattern observed.

In the follow-up analysis, we focused only on the investigation of the novel verbs to determine whether English novel words that are phonotactically legal in Czech are processed with a different level of accuracy than English novel words that are phonotactically illegal in

Czech and whether language proficiency influences the performance in any way. As in previous analyses, we started with a two-way interaction of score and set (A or B based on the novel word's phonotactic il/legality in Czech). The overview of the generalized linear model can be seen in Appendix 3. To get a simpler overview of the significant variables, an analysis of variance (ANOVA) was then run on the model. The analysis showed that none of the variables turned out significant, neither did their interaction, as seen in the table below:

	Df	Sum Sq	Mean Sq	F value	P-value
Set	1	1.618	1.618	1.618	0.224
Score	1	0.158	0.158	0.158	0.697
Set:Score	1	1.922	1.922	1.922	0.187

Table 40. The analysis of variance (ANOVA) table for the initial accuracy analysis.

Those results suggest that the participants of all proficiencies reacted with a similar level of accuracy to both sets' novel words and therefore made no difference between novel words phonotactically legal or illegal in Czech. As in the previous analysis on reaction times, it would seem that Czech functions neither as a facilitator nor as an inhibitor, not affecting the accuracy of the performance. However, also in this case, when mean accuracy was counted for the A and B subset for each participant without random effects and an independent samples t-test was run between those two subsets (mean A and B accuracy), its results turned out significant: $t(29) = 3.72$, $p < .001$, two-tailed. As in the previous interference analysis with reaction times, these findings suggest that some interference effects might be at play during L2 lexical access, though they may be too weak to reach significance in a linear-mixed model. Plotting was needed to understand whether Czech functioned as a facilitator or an inhibitor of accuracy:

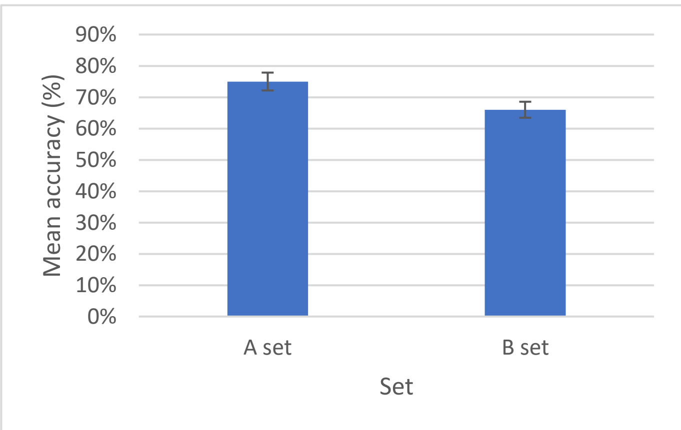

Figure 26. Mean accuracy for the A and B sets.

The plot shows that the participants were more accurate in absolute value when presented with the novel words of the A set (that are phonotactically legal in Czech) than with those of the B set (that is phonotactically illegal in Czech). Similar to the previous analysis, those findings seem to indicate that Czech functions as a facilitator of performance rather than as an inhibitor, speeding up reaction times and promoting higher accuracy, though this interference appears to be relatively weak. However, the results of the independent samples t-test did not take random effects into consideration; we must be therefore careful when interpreting the results. Nonetheless, Czech seems to somewhat affect the performance also in terms of accuracy, facilitating the reaction.

5.5 Discussion

5.5.1 Summary of the findings

The reaction-times analysis has shown the participants took longer to discriminate the novel words than the real words. The subsequent analysis of the reaction to the novel-word stimuli suggested that Czech might have had a weak facilitatory effect on the reaction times since the participants reacted more quickly in absolute value to the A-set novel-word stimuli (that are phonotactically legal in Czech).

The accuracy analysis has shown that the participants were similarly accurate when presented with the real and novel words, making no difference between the two stimulus types or performing differently among varying language scores. Further analysis of the interaction between the word type and the language score has shown that the score significantly influences accuracy only in real-word stimuli, with the progressive development of accuracy with raising proficiency. The subsequent analysis of the accuracy of the novel-word stimuli has suggested that, as in the reaction-times analysis, some transfer effects might be at play. Graphical visualization showed that the participants were more accurate when presented with novel words of the A set (that are phonotactically legal in Czech), indicating that Czech may function as a facilitator of performance rather than as an inhibitor, promoting higher accuracy.

5.5.2 Discussion and conclusion

Recent lexical-decision studies have shown that the phonological similarity between L1 and L2 leads to the facilitation of reaction to L2 stimuli (e.g., Jared & Kroll, 2001; or Marian et al., 2008). Studies on the effect of proficiency have also shown that the language level is a decisive factor in the performance and the organization of the two lexica, in a way that the lower language levels might be struggling with a greater effect of L1 interference (e.g., Talamas et al., 1999; Jared & Kroll, 2001; Kroll et al., 2002; or Harrington, 2006). The experiment aimed to investigate possible interference effects from L1 that might be at play not only during L2 production but also during L2 perception. Taking advantage of a language-proficiency scale, this experiment has also attempted to look at the potential development of such language transfer.

Our data suggest that Czech functions as a mild facilitator of performance, speeding up reaction times to presented stimuli and promoting higher accuracy, as observed in the absolute values of performance and post-hoc analyses. Consistently with the findings of Experiment 2, which suggested that lower-proficiency L2 learners use Czech as a facilitator of their production performance, with higher-proficiency levels and native speakers paying no apparent attention to language transfer, results of Experiment 3 suggest that Czech has facilitatory effects even in L2 perception. In the accuracy analysis of this experiment, proficiency significantly influenced accuracy only in real-word stimuli but not in novel words, with more proficient participants reacting more accurately to the presented real words. These findings can probably be explained by greater familiarity with the language and a potentially larger lexicon in advanced speakers. However, proficiency appears to be irrelevant for the recognition of novel verbs. Here, all proficiency scores showed mild facilitation for novel words phonotactically legal in both English and Czech. This development (or the lack of it) is rather interesting. A mild facilitatory effect of L1 was found not only at the lowest language levels but also in more proficient L2 speakers. Thus, the mother tongue seems to facilitate (to some small extent) the lexical recognition both in reaction times and accuracy at all levels.

These results open the door to a more embracing view of L1 in L2 classrooms since some effect of the mother tongue seems to be present even in more proficient learners. Referring back to the theories of language (in)dependence and (non)selectiveness, the findings of this experiment seem to speak in favour of two dependent (integrated using van Heuven et al.'s terminology) lexica, even though the result is certainly not clear-cut, and the interference reported is relatively mild.

5.5.3 Methodological outcomes

These mild results may be a consequence of methodological limitations, and further research may offer neater findings. As stated above in the reaction-times analysis, a significant number of trials had to be removed when measuring response speed because the participants responded incorrectly to the presented stimuli (by identifying a novel-word stimulus as a real English word or vice versa) – see also the rather low percentage of overall accuracy in the descriptive statistics of the accuracy analysis. This has resulted in removing some of the data from the analysis (since we were interested in correct answers only for the reaction times analysis), and it considerably reduced the size of the reaction times dataset. Lexical decision tasks are frequently used with native speakers, who tend to be at ceiling and thus offer "complete" reaction times datasets. However, with our L2 speakers, the similarity to existing English words seemed to be too prominent with several words, resulting in their incorrect discrimination, particularly with lower proficiencies (shown by the significant proficiency effect on real-word accuracy). On top of that, previous lexical decision tasks performed by L2 learners have used the novel words as a linguistic control condition, focusing mainly on performance on the real words. Our focus here was reversed: we examined the reaction times and accuracy to the novel words, and the lowered accuracy towards this stimuli type was, therefore, unwelcomed. For these reasons, it seems to be methodologically advisable to carefully control for analogy effects in the future to lower these effects as much as possible, i.e., to pay careful attention to how phonetically similar the novel words are to the existing English words, and possibly involve this level of similarity as one of the control conditions in the model.

5.5.4 Implications

Despite these limitations, the experiment still offered new insights into lexical access in second language learners, spanning from proficiency effects in real word recognition to the mild transfer effects observed in reaction times and accuracy. This is, to our knowledge, the first study investigating these effects in Czech learners of English, and this data thus complements previous work on other languages that has shown similar interference effects (Jared & Kroll, 2001; Lemhöfer & Dijkstra, 2004; Harrington, 2006; Marian et al., 2008; Brenders et al., 2011,). The language-proficiency scale also allowed us to look at the potential development of such language transfer (that, to our surprise, language development did not matter, and language transfer did not change across various levels of proficiency).

The results of our experiment have brought two significant implications. One of them is methodological – it seems to be methodologically advisable to carefully control for the similarity to existing words when presenting this type of task to L2 learners, to lower those effects as much as possible. The second implication is pedagogical since this experiment allowed us to look deeper into processes underlying language transfer behind L2 perception and consider its application in teaching practice. Since Czech seems to have a mild facilitatory effect on L2 learning, it may be used actively in an L2 classroom and in the preparation of learning materials. Taking the learner's L1 into account may be beneficial for the L2 learner and the quality and speed of their learning.

In conclusion, this experiment has allowed us to look into the mechanisms that underlie the interplay of L1 and L2 in an auditory lexical decision task in L2 learners on a proficiency scale. The findings have shown that a mild facilitatory effect of L1 can be found not only at the lowest language levels but also in more proficient L2 speakers (as evidence by the constancy of the effect across the whole proficiency scale). These results open the door to a more embracing view of L1 in L2 classrooms since the effect of the mother tongue on the L2 seems to be ever-present.

6. GENERAL DISCUSSION AND CONCLUSION

This dissertation aimed to explore the development of inflectional morphology in L2 learners of English, both in terms of how inflectional morphology is aurally perceived and produced (with additional focus on how L1 Czech influences such performance). Our primary focus was on sublexical development (i.e., no access to the lexicon) and, for this purpose, novel words were employed in all experiments.

Previous studies have shown that L2 learners face specific challenges when processing and producing L2 inflectional morphology. Several possible theories have been proposed as to where the difference and the source of the struggle might lie, yet no clear answer has been given so far. However, the learner's L1 seems to be influencing their morphological behaviour and, therefore, needs to be taken into consideration when mapping L2 inflectional morphology. In order to test the development of inflectional morphology in L2 learners and the influence of Czech, we performed three experiments with three different methodologies that would best fit our purposes and offer a complex picture of this development. A progression of language levels or proficiency scores was used in all cases to allow us to properly map the development of inflectional morphology in L2 learners and describe its specifics.

6.1 Brief summary of the experiments

The first experiment focused on the perception of word-final bound morphemes. Using novel words with and without potential morphosyntactic information and a phonetic control condition, we attempted to find evidence either for stems being stored separately from inflectional morphemes (that would manifest itself by morpheme stripping) or for analogical processing (that would manifest in a phonotactic probability effect) in novel morphologically-inflected forms with no access to the lexicon. Previous work by Post et al. (2008) and Cilibrasi (2016) show evidence for native speakers performing morpheme stripping even at the sublexical level, with the parsing being synergistic to the analysis of the stem. Previous L2 studies have shown support for both online decomposition of inflectional morphology in L2 learners (e.g., Hahne et al., 2003; Voga et al., 2014; Kimppa et al., 2019) and for its lack (e.g., Silva & Clahsen, 2008; Clahsen & Neubauer, 2010; or Clashen, 2010). Our previous findings

(Jiránková, 2017) show that more proficient English L2 learners with Czech as L1 might be sensitive to the presence of potential morphosyntactic information applied to novel words. The question of how lower-level participants process English inflections, however, remained unanswered. Given the difference in inflectional richness between Czech and English, we hypothesized that either our participants would use morphological decomposition only at later stages of L2 learning, or their L1 will influence their performance and they will be more sensitive to morphological processes even in English, performing morpheme stripping in a similar way to native speakers from early stages. Our hypotheses were tested with a perception experiment, in which the participants were supposed to decide whether the items of the minimal pair they heard were identical or different. Reaction times and pressed keys were recorded for further analyses, and the effect of Czech was investigated by the inclusion of L2 learners with low proficiency levels of English.

The first experiment has allowed us to look more closely into the mechanisms underlying the perception of inflectional morphology in L2 learners of English, showing that morphological decomposition might be accompanied by frequency and length effects in the sublexicon, and a more embracing explanation of such perception is needed. Nonetheless, we presented further evidence that inflectional morphemes can be detected sublexically and that, similarly to monolingual English speakers, L2 learners of English are likely sensitive to morphosyntactic information in novel words. Our study has also shown that L2 participants might analyse the phonological quality of the bound morpheme together with the quality of the stem (in line with Jiránková, 2017), and it demonstrated that the perception of English inflectional morphology might be possibly explained in terms of the decompositional model, with parallel frequency effects. In summary, our results contribute to the debate of how morphologically-inflected forms are stored, suggesting that both morphological decomposition and analogical processing might be active in the sublexicon for L2 speakers. Interestingly enough, English proficiency played only a small role in this experiment: The pattern of L2 processing of inflectional morphology was comparable to that of native speakers in all language levels. Since the same pattern effect was observed at the lowest proficiency levels with little knowledge of English and other English proficiency levels, it is likely that L1 Czech was a decisive factor in the performance: there might be a general sensitivity of Czech learners to inflectional morphology, given Czech morphological richness.

The second experiment focused on active spoken production of novel verbs with different similarities to existing regular or irregular English verbs. Using an elicitation task, we aimed to find evidence either for the dual-route model, which posits that regular past-tense forms are produced by default rules, while irregular past-tense forms are produced analogically to stored examples, or for the single-route model, which assumes that both all past-tense forms are produced by a single mechanism, i.e., by its similarity to existing regular and irregular verbs. We have, therefore, attempted to search for evidence either in support of the use of rules or analogy with regular past-tense inflection in L2 learners of English at different proficiency stages. Previous studies by Albright and Hayes (2003), Ambridge (2010), and Blything et al. (2018) provide evidence that both adult and child native speakers of English base the form of the verb's past tense on the novel verb's similarity to existing regular verbs, producing more regular past-tense forms with words highly similar to existing regulars. Native speakers, therefore, prefer the use of analogy to the use of default past-tense rules. Previous L2 studies have shown that the dual-route model might not be capable of fully explaining and describing the production of L2 learners, showing how word frequency can affect the production of regular verbs (e.g., Beck, 1997 or Ellis & Schmidt, 1998). Previous research also showed the decisive role of the novel verb's phonological resemblance to existing (ir)regulars on the production of past-tense forms (e.g., Murphy, 2000, Agathopoulou, 2009, or Cuskley et al., 2015). Based on these findings, we hypothesized that the dual-route model would not be able to fully describe our data, and an alternative production mechanism will be used instead by L2 learners. Our hypotheses were tested with an elicitation task in which the participants were prompted to finish past-tense sentences using a presented novel verb in its past form. Produced forms and reaction times were recorded for further analyses.

The second experiment enabled us to look more closely into the processes underlying morphological productivity in L2 learners of English, and the results showed a gradual progression from the application of default rules to the application of analogical patterns spotted during the process of language learning. The data suggest that the lowest proficiency groups (A1, A2, and B1) generate regular past forms by the default rule of adding *-ed* to the novel verb's stem, seemingly irrespective of the novel verb's similarity to existing regular verbs. In contrast, the more proficient groups (B2 and C1) and native speakers of English seem to have developed a morphological system in which the production of regular past forms depends on the novel verb's phonological similarity to existing verbs based on analogy (in line with previous work on L2 learners of English – e.g., Beck, 1997; Ellis & Schmidt, 1997; Murphy, 2000; Agathopoulou, 2009; Cuskley et al., 2015). For the native speakers, it was the novel verb's similarity to existing regular verbs that played a significant role in the production of past forms, in line with the single-route model and previous findings by Albright and Hayes (2003), Ambridge (2010), and Blything et al. (2018). The second-language speakers thus come closer to adult native speakers with higher proficiency of B2 and C1 levels. The analysis of reaction times has also shown that for the lowest language levels (A1 and A2), Czech implicitly functions as a facilitator in their performance: Novel words that were legal in both English and Czech were processed more quickly. The qualitative analysis has also uncovered some specific pronunciation issues, such as the use of the full pronunciation of bound morphemes.

The third experiment employed a lexical decision task and the novel verbs from Experiment 2 (that were divided into phonotactically legal and illegal in Czech) to look more closely at potential transfer effects from L1 Czech. Recent findings show that L1 does not seem to be easily separated from the processing in L2 (Kroll et al., 2002; Costa & Santesteban, 2004; Sunderman & Kroll, 2006; Schoonbaert et al., 2009; Macizo et al., 2010). Research on interference effects generally makes a distinction between the word association model, with a direct mediation of L1, and the concept mediation model, with no mediation of L1 (Potter et al., 1984). Recent studies show that L2 learners decrease their dependence on L1 with higher proficiency and access L2 words directly (Kroll et al., 2002). In connection to the relation between the two lexica and their storage, van Heuven et al. (1998) proposed a distinction between "language-dependent lexica" and "language-independent lexica." Concerning the online processing of words from the two lexica, a difference is often made between selective or non-selective access. Several research findings show that during language processing, L2 learners seem to active the lexica of both languages in non-selective language access (Van Heuven et al., 1998; Lemhöfer & Dijkstra, 2004; Brenders et al., 2011). Previous studies have also shown that proficiency is a decisive factor in the organization of the two lexica with greater interference for lower language levels (see, for instance, Kroll, 1993). This led us to investigate the effect of L1 on English novel words in L2 learners of English at different proficiency stages. We hypothesized that Czech might function either as a facilitator of the performance (with novel verbs that were phonotactically legal in Czech promoting quicker reaction times due to greater phonetic familiarity) or as its inhibitor (with the same novel verbs promoting longer reaction times due to the complexity of using both languages simultaneously). Our hypotheses were tested with a lexical decision task employing real English verbs and novel verbs deemed phonotactically legal in English. The participants were asked to decide if a presented verb was a real English word or not. Reaction times and pressed keys were recorded for further analyses.

The third experiment has allowed us to look into the mechanisms underlying the interplay of L1 and L2 in an auditory lexical decision task in L2 learners, and its findings have shown that a mild facilitatory effect of L1 can be found not only at the lowest language levels but also in more proficient L2 speakers. The subsequent analysis of the accuracy of the novel-word stimuli suggested that, as in the reaction-times analysis, some interference effects might be at play as well, indicating that Czech may function as a facilitator of performance rather than as an inhibitor, promoting higher accuracy. These data thus complement previous work on other languages that have shown similar transfer effects (Lemhöfer & Dijkstra, 2004; Brenders et al., 2011, Marian et al., 2008, Jared and Kroll, 2001, Harrington 2006). These results also open the door to a more embracing view of L1 in L2 classrooms since the effect of the mother tongue seems to be ever-present and may thus not be ignored also when dealing with proficient learners.

6.2 General discussion

All in all, the three experiments have enabled us to look more closely into the perception and production of inflectional morphology in L2 learners of English and simultaneously comment on its storage, online processing, and the effect of L1. Moreover, having recruited participants at different language levels, we were also able to look at morphological development from the initial stages of L2 acquisition to near-native proficiency.

The findings of the perception experiment extend previous research on morphological processing in novel words (e.g., Post et al., 2008; Cilibrasi, 2016; Jiránková, 2017; Cilibrasi et al., 2019), providing further evidence that inflectional morphemes can be detected sublexically. The data also suggest that not only monolingual English speakers (as tested by Post et al., 2008; Cilibrasi, 2016; or Cilibrasi et al., 2019), but also L2 speakers of English are sensitive to the presence of morphosyntactic information in novel words. The results have also confirmed that L2 learners might implicitly analyse the phonological quality of the stem and the affix simultaneously, in line with previous research done, for instance, by Pater (2006), Harris (2011), Cilibrasi (2016), or that of Grainger and Ziegler (2011) who proposed the existence of "a fine-grained parser" capable of detecting specific grapheme groups that may carry morphological information. Contrary to previous studies on monolingual speakers and in line with my MA thesis on more proficient L2 learners of English (Jiránková, 2017), this thesis has also shown that detecting the presence of morphological information is not the only strategy active in L2 learners – parallel analogical processing of novel words based on their phonotactic probabilities was also found. In this experiment, L2 learners thus seemed to use a mix of both decompositional and atomist strategies in a dual-route mechanism since in their perception of regular morphologically-inflected novel words, they used both decomposition (in line with the decomposition theory, which posits that regular word forms are produced by the application of a default rule and perceived as such; see, for instance, Berko, 1958; Pinker & Ullman, 2002; Post et al., 2008) and analogical processing (in line with the atomist theory based on phonotactic probabilities, which posits that regular word forms are produced based on their phonological similarities to existing words and, during perception, higher phonotactic probabilities amount to quicker reaction times; see, for instance, Bertram et al., 2000; Stemberger & MacWhinney, 1986; Tomasello, 2000). Our findings also support previous L2 research that proposed that native speakers and L2 learners show a similar amount of sensitivity to the morphological structure of a word, performing decomposition in language perception (e.g., Basnight-Brown et al., 2007; Feldman et al., 2010; Coughlin & Tremblay, 2015; Coughlin et al., 2019). Contrary to some views (e.g., Ullman, 2005), we have also managed to show that this sensitivity does not seem to be influenced by L2 proficiency (see, for instance, Coughlin et al., 2019). This finding might be caused by the influence of a morphologically rich L1, as evidence by previous research on L1 transfer in L2 perception (e.g., Juffs, 1998; White, 2003; Bliss, 2006; Basnight-Brown et al., 2007; Chen et al., 2007; Feldman et al., 2010) that argues that the learner's sensitivity to morphosyntactic information might be affected by their L1.

The production experiment extends recent production tasks using novel words (e.g., Albright & Hayes, 2003; Ambridge, 2010; or Blything et al., 2018) that provided evidence for the use of the analogical single-route model in language production. Similarly to native English-speaking children and adults (Albright & Hayes, 2003; Ambridge, 2010; Blything et al., 2018) and L2 learners of English (Beck, 1997; Ellis & Schmidt, 1997; Murphy, 2000; Agathopoulou, 2009; Cuskley et al., 2015), our findings also show that the novel word's similarity to existing English verbs seems to be playing a decisive role in the production of English past tense at the B2 and C1 levels and in native speakers, in line with the single-route model (e.g., Bybee, 1995). Contrary to that, the A1, A2, and B1 participants applied default past rules in their production, much in line with previous research supporting the dual-route mechanism (Prasada & Pinker, 1993; Chialant & Caramazza, 1995; Clahsen, 1999). Our findings also support previous L2 studies, which suggest that the dual-route model is not fully capable of describing the production of L2 learners (e.g., Beck, 1997; Ellis & Schmidt, 1998; Murphy, 2000;

Agathopoulou, 2009; Cuskley et al., 2015). In our experiment, the predictions of the dual-route model were found only in the lower-proficiency levels.

The two experiments showed important differences in relation to proficiency groups: While the perception experiment showed a difference in overall accuracy between levels A1 + A2 and the rest of the proficiencies, but the same pattern of performance in all groups, the production experiment showed a *qualitative* difference between levels A1 to B1 and the rest of the proficiency groups. In the perception experiment, all L2 learners seem to perform similarly (with similar reaction times differences to individual conditions) to the English native speakers that had been previously tested by Cilibrasi (2016); on the contrary, in the production experiment, only the B2 and C1 levels use analogy similarly to native speakers. The lower proficiency groups relied on a qualitatively different system to form new inflected verbs and applied rule-like processing.

In terms of processing models, an interesting distinction thus emerges between L2 perception and production. In perception, the L2 performance showed a mix of analogical and decompositional strategies in a dual route mechanism observable in all groups. In production, the proficiency groups demonstrated a transition from a dual-route system (at A1 to B1 levels) to a single-route system (at B2 and C1 levels and native speakers). The production development might be likely explained by growing language input and learners' heightened awareness of analogical patterns in English. The initial application of rules (as the easiest and most automatic strategy) is slowly replaced by the application of analogical patterns to new forms. In comparison, the use of both systems in the perception experiment is very peculiar. Morpheme decomposition might be possibly explained by the higher sensitivity of all L2 participants to word-final bound morphemes since Czech is a highly inflectional language and paying attention to word-endings is necessary for properly understanding the meaning of a sentence (see also Ku & Anderson, 2003). The use of analogy, however, indicates that already at the A1 level, participants have to have some general (and very likely implicit) knowledge of English phonotactics and what is and what is not phonotactically legal and probable in their L2.

Theories of how inflectional morphology is processed in real time usually make a strict distinction between the effects of the single- and dual-route model, trying to find evidence either for the former or for the latter. However, in our research, we have attempted to offer a less diving conclusion, very much in line with the so-called "redundant models" (e.g., Schreuder et al., 1999). These models assume that the two types of analyses (i.e., analogy and rule application) operate in parallel and in a complementary way, but each system has varying importance connected to how frequent the presented item is in the language: while frequent items (or, when speaking of novel words, items with high phonotactic probability) are to be decoded and produced with the rote-based system, less frequent items (or, when speaking of novel words, items with low phonotactic probability) are to be decoded and produced with the rule-based system (Cilibrasi et al., 2019). Using Schreuder et al.'s proposal, we can justify the simultaneous activation of morphological decomposition and frequency effects in Experiment 1 and a progressive transition from the application of rules to the use of analogy in L2 participants tested in Experiment 2. Even though this explanation of our seemingly contrary findings might be inconsistent with some influential views on language processing (e.g., Chomsky's principle of economy, see Chomsky, 1995), the choice between single- vs. dual-route model does not have to be exclusively one or the other. The most descriptive answer enveloping all of our research findings might be simply that both mechanisms work in parallel, and each of them serves its own purposes.

The lexical decision experiment extends previous research on cross-language transfer effects. In line with recent findings (e.g., Kroll et al., 2002; Lemhöfer & Dijkstra, 2004; Harrington, 2006; Marian et al., 2008; Brenders et al., 2011), it showed an effect of L1 on L2 processing. Contrary to some studies on the effect of proficiency and L1-L2 interaction, our findings do not show that lower language levels might be struggling with a greater effect of L1 interference (e.g., Talamas et al., 1999; Jared & Kroll, 2001; Kroll et al., 2002; or Harrington, 2006): Our results show that Czech has a minor facilitatory effect not only at the lowest proficiency levels but also in more proficient L2 learners. Our findings thus do not seem to support the view that the dependence on L1 decreases with higher proficiency (see, for instance, Kroll et al., 2002); it seems instead that L1 mildly facilitates L2 performance in the whole course of second-language acquisition. Regarding the relation between the two lexica and their storage, our data seem to speak in favour of an integrated lexicon for both activated languages. In terms of access, our data provide further evidence for the presence of non-selective language access (see, for instance, Van Heuven et al., 1998; Lemhöfer & Dijkstra, 2004; Brenders et al., 2011), in which both languages are activated simultaneously. In general, our findings on language transfer and the effect of L1 in all three experiments have shown that Czech might have a facilitatory effect both on the perception and production of English inflectional morphology. This might be caused by the fact that Czech is morphologically much richer than

English and its native speakers are, therefore, generally more aware of inflectional morphology, not only in their mother tongue but also in other languages they learn.

6.3 Critical evaluation of the methods and stimuli

Similarly to any other research project, when preparing this study, we had to make difficult decisions that were not always fully satisfactory, and potential problems were handled with the best intention in mind. It is, however, essential to be aware of these critical issues when evaluating the findings of this thesis.

First, the whole thesis is focused on testing L2 learners of English at various stages of proficiency. Even though the grouping used in Experiments 1 and 2 differs from that used in Experiment 3 (where we used proficiency scale instead of categorical language levels), the proficiencies were all assigned on the basis of a placement test given to the students of the Faculty language school. The major weakness of this placement test is that it tests only grammar and vocabulary, excluding other language skills (i.e., reading, writing, and speaking). The "attained language level" is, therefore, only an estimate of the student's real level in all these domains. Needless to say, the outcomes of these placement tests have been verified by in-person discussions with a random sample of the test takers[50], and the final language levels correlated with each other. This is the main reason why we decided to keep these proficiency scores even though the placement test was mainly lexico-grammatical.

Another potential issue that relates to all three experiments is the monosyllabicity of the presented novel word stimuli as opposed to the typical multisyllabicity of Czech words (thus, even if the novel word is phonotactically legal in Czech, it deviates from Czech in terms of its length to some degree). One of the legitimate critiques of the stimuli presented to the L2 learners in order to investigate language transfer from L1 is precisely the length of the stimuli. This difference is caused by the different typology of the two languages: while Czech is an inflectional language and as such tends to express (not only) its grammatical categories by adding suffixes and creates multisyllabic words, English is an analytic language that tends to cumulate monosyllabic words and express potential grammatical categories not only by adding

[50] Upon administering the written placement test, the author of the test chose a random sample of particpants and invited them for an interview. There, he attempted to measure whether their oral production corresponded to the score obtained in the written placement test. He found that the realibility of the final score in the written part correlated with the score that the participants received in the oral part. The written placement test was, therefore, deemed usable for further proficiency division in various experiments and learner corpora.

non-syllabic suffixes but also by the addition of other grammatical (functional) words. When investigating speech rate transfer between English and Czech, Jiránková et al. (2019) found that the difference between mean syllable-word ratio (i.e., the mean number of syllables in one syllable) between Czech and English was 0.5 syllables per word (Gráf (2015) found a mean ratio of 1.3 syllables per word in informal English speech; Jiránková et al. (2019) found a mean ratio of 1.8 syllables per word for spoken Czech). Typology, therefore, plays a role in these two languages, and L2 learners will always have to adapt typologically when using English. Working with monosyllabic novel stimuli (that correspond to predominantly monosyllabic English words) might, therefore, lead to a more authentic investigation since all experiment test participants' (even implicit) knowledge of English. However, the number of syllables per word used in this study needs to be borne in mind when interpreting the transfer findings of this study – even though the data analysis might show effects of L1, Czech stimuli would actually likely ask for polysyllabicity instead of English monosyllabicity. Our findings concerning language transfer thus need to be interpreted carefully.

Another issue regards some of the sub-sample sizes. To look at the possible effect of L1 Czech on the production of English past-tense, Experiment 1 involved a sample of 4 A0 L2 learners, i.e., learners with zero or very little knowledge of English. The findings gathered for their language group might give us a very general idea of what the situation looks like for Czech students untouched by English; however, the number is way too low to reach any specific conclusions. More participants would need to be tested; their recruitment is, however, rather problematic at a time when English is ever-present even in the Czech cultural surroundings, and adults with zero or little knowledge of English need to be tested. Similarly, the sample of native speakers in Experiment 2 is rather small, even though we can derive data from other studies using the same methodology, so this seems to be less of a problem.

Another problematic issue in Experiment 1 is the significant contrast between the non-morphological and phonological control conditions. The control condition was included in order to ensure that the contrast between the morphological and non-morphological condition is not purely phonological (since the morphological condition ended in voiced phonemes, while the non-morphological ended in their voiceless variants). The contrast between the non-morphological and control condition cannot be explained morphologically (since the control condition does not include any potential morphosyntactic information), but it may be due to phonology (since the non-morphological condition ends in a voiceless phoneme and the control condition in a voiced one). This explanation could easily undermine our interpretation of the significant contrast between the morphological and the non-morphological condition since the same logic might be applied to this contrast as well (and the control condition was added to avoid this explanation in the first place). As of now, it is not completely possible to say what kind of impact morphology and phonology have on this contrast, and further research is needed to establish it.

Another issue regards Czech phonotactic probabilities. Experiment 2 divided the stimuli adopted from Albright and Hayes (2003) and Blything et al. (2018) into two halves: 16 novel words that were deemed phonotactically legal both in English and Czech and 16 novel words that were phonotactically legal only in English but not in Czech. As a reference point, a paper by Štrum and Lukeš (2017) was used that provides all possible Czech syllable onsets and codas, but it does not give information about syllable nuclei. For the lack of a better phonotactic calculator for Czech, the stimuli had to be divided based on whether both the syllable onset and coda were deemed phonotactically legal in Czech or whether either the syllable onset or coda or both parts were deemed phonotactically illegal in Czech, omitting the syllable nuclei. Our stimuli division, therefore, could not provide a full and complex picture of Czech phonotactic (il)legality. However, given the lack of any automatized calculator that would allow stimuli importation and export the word's similarity or dissimilarity to existing Czech words, this phonotactic analysis was the only relevant reference point that could be used for the division. Even though the study omits syllable nuclei, it is, at the moment, the only suitable reference point for the inclusion of Czech transfer effects available. Any assumptions on the effect of Czech, however, need to be read with this in mind. Another potential drawback related to this is the lack of any automatized counter of phonotactic probabilities for Czech that would be very useful in all three experiments. The first experiment controls for English phonotactic probabilities, but the other two experiments involve only English and Czech phonotactic legality. Since phonotactic probabilities are very important in morphological processing but no automatized counter is available for Czech, phonological neighborhood density could be used in further research and calculated in such programs as *Clearpond*.

The qualitative analysis in Experiment 2 has also uncovered several instances of potential priming. Those were evident in the section focused on the full pronunciation of past-tense bound morphemes, e.g., *dized* pronounced fully, possibly as an analogical repetition of the form *dizes* that preceded the prompted past-tense form, or in the section focused on the insertion of *-s-* phonemes where the elicited past-tense form (e.g., *panksed*) could be, once again, influenced by the preceding present-simple form *panks*. These word forms and the possible priming were unintentional, and the findings in those two categories have to be also understood with the priming effect in mind.

The last issue is related to the lexical decision task used in Experiment 3. As reported, we had to remove a significant number of trials from the reaction-times data due to incorrect responses to the presented stimuli (when participants identified a novel-word stimulus as a real English word or vice versa), which reduced the dataset size considerably. The similarity to existing English words seemed to be too prominent with several novel words, especially in lower proficiencies. Given that we were mainly interested in the reaction times and accuracy to the novel words, the lowered accuracy was unwelcomed. For those reasons, it seems to be methodologically advisable to carefully control for the novel word's similarity to existing words in future research (e.g., by including its similarity score as a control variable) to lower these effects as much as possible.

This critical reflection of our experiments, therefore, advises that the findings were interpreted with caution and understood in terms of these limitations. The author did her best to reduce the downsides of the studies and offer methodologically precise procedures; however, several issues made themselves evident only after the testing process. Nevertheless, these research limitations offer valuable ground for further research, new methodological improvements, and refined replications.

6.4 Directions for further research

This study investigated the development of the production and perception of inflectional morphology in L2 learners of English with Czech as L1. Since Czech is an inflectionally rich language, it creates an interesting contrast with the notoriously poor inflectional system of English and provides a suitable ground for investigating possible transfer effects of Czech on the acquisition of English inflectional morphology. A certain effect of L1 has been proved in all three experiments, although with varying significance. The inclusion of lower proficiency participants in Experiment 1 has suggested that the morphological awareness of Czech L2 learners might be caused by their strengthened sensitivity to word-final morphology in Czech. In this respect, more A0 L2 learners should be recruited to confirm or refute this finding.

In contrast, recruiting participants with a morphologically-poorer L1 than Czech and comparable with English could be an interesting extension of the present research. Native speakers of Asian languages such as Thai or Chinese would be suitable candidates since those speakers usually struggle with English inflectional morphology even in repetition tasks. The difference in writing systems could, however, turn out problematic, and the study would need to focus on aural perception and spoken production only. This would give us a more detailed view of the L2 perception and production of English inflectional morphology with an L1 that is comparably poor in terms of morphology.

6.5 Some ideas for a practical application of my findings

The findings of this thesis are also applicable in teaching practice since they offer direct pedagogical implications. One crucial question to consider is to what extent the acquisition of L2 morphology is affected by teaching and to what extent it is the result of some built-in syllabus (i.e., the learner's internal language representations; see, for instance, Corder, 1967). In my experience, inflectional morphology is taught very explicitly at the lowest levels, at least in the Czech Republic, where the experiments were conducted. Our results show that with rising proficiency, learners are more prone to noticing analogy patterns and applying them to new items. One could argue that even the explicit teaching should follow the natural order of acquisition and not violate it (for instance, by introducing the unproductive *-en* morpheme for plurals before a more productive *-s*). Moreover, it has also been observed in language classrooms before that L2 students do not necessarily learn all the grammar they are taught and might instead internalize only those features that they are ready to learn and use. The teaching and students' build-in syllabi should thus go hand in hand. We will now list specific outcomes that call for further didactic response and some specific questions that are closely tied to them:

1. During perception, L2 students (at all language levels) seem to be able to distinguish regular forms with potential morphosyntactic information (where the phonetic realization of the bound morpheme respects basic rules of English morphosyntax) from forms with no such information. An assumption can, therefore, be made that when L2 learners are able to hear the different phonemic realizations of the regular bound morpheme and be aware of these realizations that follow the rules of English morphophonology, they might also be able to produce these allomorphs (a generally accepted idea of proper perception preceding successful production, as described, for instance, in Antolík, 2020). So far, different realizations of bound-morpheme allomorphs are not frequently taught to Czech L2 learners of English since the learning

process involves basic phonetic instruction. Incorrect pronunciation of word-final allomorphs is, however, one of the specifics of Czech English pronunciation that keeps even proficient speakers from obtaining native-like pronunciation. Should these allomorphs be explicitly or implicitly taught then?

2. During production, L2 students start with the application of a general past-tense rule (A1-B1) and move on to the use of analogy (B2 and C1) based on the novel word's similarity to existing (ir)regular verbs, while native speakers use analogy only. More proficient language learners, therefore, eventually start using the same strategy as adult native speakers of English. Should the use of analogy for past-tense forms be supported right from the early stages of L2 acquisition?

3. The production experiment has shown that at the lowest levels (A1 and A2), Czech functions as a facilitator of the performance, while the lexical decision task has shown mild facilitatory effects of L1 Czech on the whole proficiency scale. Should Czech be actively made use of in the teaching process? Should learning initially focus on words that are phonotactically legal also in Czech?

These questions led us to a proposal of a new strategy for teaching the English past tense, in which individual language levels would learn verbs and their past-tense forms suitable for their language level using analogy. In order to make even lower-proficiency learners reach the stage of adult native speakers earlier, the use of analogy in past-tense teaching would be promoted from the very start of L2 learning since L2 learners eventually start using the same strategy as adult native speakers. This way, the whole learning process can be sped up, and analogy can be promoted from the very beginning (this will also lead to greater language awareness). Our data did not show the use of analogy with lower-level students (A1 to B1) in production; however, Federici et al. (1996) show how neglected the use of analogy for language learning has been so far. Investigating the advantages of the use of analogy in language learning, Gentner and Namy (2006) showed that by comparing two concepts, students could realize the similarities between them, which then promotes easier acquisition. Moreover, previous studies show that the use of analogy for past-tense learning can be beneficial even for lower-proficiency learners. Yoke and Hasan (2014), for instance, showed that the use of analogy was an effective tool for promoting the acquisition of English past tense in lower-proficiency students. Therefore, there seem to be several advantages to the proposed strategy: First, L2 learners at individual language levels will learn only those (ir)regular verbs that are relevant to their level. Second, verbs will not be taught as alphabetical lists but based on structural analogical similarities between all existing verbs. Third, analogical groups will be used not only for irregular verbs (a strategy we saw used, for instance, in Quirk et al., 1985) but also for regular verbs. Fourth, for regular verbs, inductive learning of final allomorphs will be promoted. Last, but not the least, the facilitatory effect of Czech can be taken into account as well.

Having explained the main concept behind this new teaching strategy, we will now explain how we proceeded to get the desired teaching material.

1. All verbs that fall under the knowledge of an L2 learner at a given language level were imported from the English Vocabulary Profile,[51] a free online platform that enlists all vocabulary known by L2 learners at specific language levels. We imported only the British English version since we assume that most of our L2 learners will have experience with RP English presented in English coursebooks. In the extension of this teaching material, the British variant can be compared with the American one; however, not much difference is expected. We have also excluded phrasal verbs, focusing on bare verbs only, since, in phrasal verbs, the past-form is still dependent on the verbal part of the phrase. What we also need to realize is that this list does not encompass the complete vocabulary knowledge of an English native speaker – even C2 learners will not have the same lexical knowledge as native speakers. Since this material is aimed at L2 learners, the imported verb list is sufficient enough for our purposes (in further research, this list may be enriched by all other verbs known by native speakers). For each language level, a list of all verbs known by that particular language level was obtained.

[51] Accessible at https://www.englishprofile.org/wordlists.

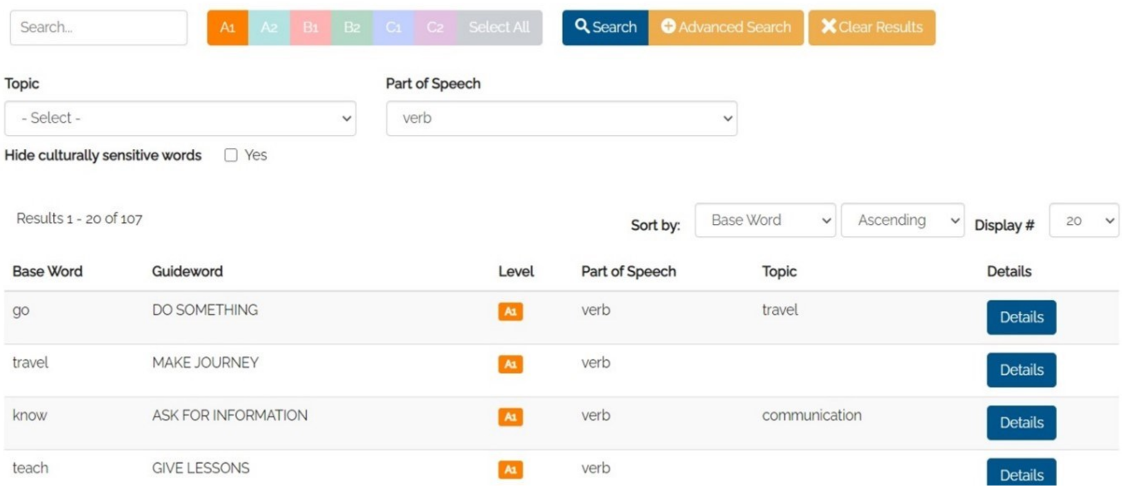

Figure 27. An example of verb singling for the A1 level in the English Vocabulary Profile.

2. As the next step, duplications needed to be erased. Several verbs appeared at different language levels with a slightly different meaning, as is natural in the process of L2 learning. However, since this material focuses solely on the verbal form and not on the nuances in the meaning, the verb in question was retained only in the list of that language level where it appeared first. We assume that the verbs from the lower levels will need to be also recycled in more proficient levels, and if an L2 learner encounters the same verb form with a slightly different meaning, the past-tense form of that form will only have to be recycled from what they already know. Only homonyms, such as *lie*, were kept in all given language levels since they change not only the meaning but also the form of their past tense.
3. The verbs in each language level were then divided into these that form the past-tense form in a regular manner and these that form past-tense forms irregularly. If both forms are possible with some words, such as *hang*, both variants are kept in the respective group.
4. Regular verbs were then divided according to their final allomorphs, i.e., if the past-tense morpheme should be pronounced as /t/, /d/, or /ɪd/ based on the quality of the stem-final phoneme.

VERB	R/IR	past tense	past participle	class
ask	R	asked	asked	t
cook	R	cooked	cooked	t
dance	R	danced	danced	t
finish	R	finished	finished	t
kick	R	kicked	kicked	t
like	R	liked	liked	t
look	R	looked	looked	t
practise	R	practised	practised	t
smoke	R	smoked	smoked	t
stop	R	stopped	stopped	t
talk	R	talked	talked	t
walk	R	walked	walked	t
wash	R	washed	washed	t
watch	R	watched	watched	t
work	R	worked	worked	t
invite	R	invited	invited	ɪd
need	R	needed	needed	ɪd
paint	R	painted	painted	ɪd
start	R	started	started	ɪd
visit	R	visited	visited	ɪd
wait	R	waited	waited	ɪd
want	R	wanted	wanted	ɪd
answer	R	answered	answered	d
carry	R	carried	carried	d
change	R	changed	changed	d
clean	R	cleaned	cleaned	d
close	R	closed	closed	d
colour	R	coloured	coloured	d
die	R	died	died	d
enjoy	R	enjoyed	enjoyed	d

Figure 28. An example of how the regular list was initially divided for the A1 level.

5. Irregular verbs were further divided into individual groups as proposed by Quirk et al. (1985) according to structural analogy among individual group members.

VERB	R/IR	past tense	past participle	class	subcategory
learn	IR	learnt	learnt	1	A
send	IR	sent	sent	1	B
have	IR	had	had	1	C
make	IR	made	made	1	C
show	IR	showed	shown	2	
feel	IR	felt	felt	3	A
leave	IR	left	left	3	A
sleep	IR	slept	slept	3	A
buy	IR	bought	bought	3	B
catch	IR	caught	caught	3	B
teach	IR	taught	taught	3	B
think	IR	thought	thought	3	B
tell	IR	told	told	3	D
hear	IR	heard	heard	3	E
say	IR	said	said	3	F
choose	IR	chose	chosen	4	Aa
speak	IR	spoke	spoken	4	Aa
wear	IR	wore	worn	4	Ab
draw	IR	drew	drawn	4	B
know	IR	knew	known	4	Ba
take	IR	took	taken	4	Bb
give	IR	gave	given	4	Bc
eat	IR	ate	eaten	4	Bf
see	IR	saw	seen	4	Bg
do	IR	did	done	4	C
ride	IR	rode	ridden	4	Ca
write	IR	wrote	written	4	Ca
fly	IR	flew	flown	4	Cb
drive	IR	drove	driven	4	Cc
meet	IR	met	met	6	A
read	IR	read	read	6	A
find	IR	found	found	6	C
sit	IR	sat	sat	6	E
get	IR	got	got/gotten	6	F

Figure 29. An example of how the irregular list was divided for the A1 level based on structural analogy.

Quirk et al. (1985: 104) classifies around 250 English irregular verbs into seven classes based on three set criteria (see the full classification in the table below):

(a) The past-tense suffix used: *-ed* (e.g., *burned*), *-t* (e.g., *dreamt*), *-n* (e.g., *torn*)
(b) The identity or difference of the past-tense form and the past participle: e.g., *met- met*, *sawed – sawed/sawn*
(c) The quality of the nuclei vowel based on the identity or difference with the base vowel: e.g., *cut – cut - cut*, *swim – swam - swum*

CLASS	USE OF SUFFIX	V-*ed* IDENTITY	VOWEL IDENTITY	*Example*		
				V	V-*ed*$_1$	V-*ed*$_2$
1	+	+	+	*burn*	*burned/burnt*	*burned/burnt*
2	+	±	+	*saw*	*sawed*	*sawed/sawn*
3	+	+	−	*bring*	*brought*	*brought*
4	+	−	−	*break*	*broke*	*broken*
5	−	+	+	*cut*	*cut*	*cut*
6	−	+	−	*strike*	*struck*	*struck*
7	−	−	−	*swim*	*swam*	*swum*

Table 41. Quirk et al. (1985: 104)' classification of irregular verbs.

The + symbol means that the relevant criterion is present in the given class, the – symbol indicates the absence of such a criterion. If both symbols appear simultaneously, variability is present in the class. The table is further divided into several subcategories based on the phonological changes that the base undergoes to create past-tense forms and past participles (PP in the table) (the full table can be seen below).

Class	Subcategory	Use of suffix	V-*ed* identity	Vowel quality	Example
1	A	-t/-ed	+	+	*burn, learn*
	B	-d → -t	+	+	*build, send*
	C	→ -d	+	+	*have, make*
2		+	±	+	*shave, show*
3	A	+	+	/iː/ → /ɛ/	*dream, leave*
	B	+	+	→ /ɔː/	*bring, buy*
	C	+	+	/uː/ → /ɒ/	*lose*

	D	+	+	/ɛ/ → /əʊ/	*sell, tell*
	E	+	+	/iːʳ/ → /ɜːʳ/	*hear*
	F	+	+	/eɪ/ → /ɛ/	*say*
4	A	base form ≠ past forms vowel			
	Aa	+	-	→ /əʊ/	*break, choose*
	Ab	+	-	/eɪəʳ/ → /ɔːʳ/	*swear, wear*
	Ac	+	-	/aɪ/ → /ɪ/	*bite, hide*
	Ad	+	-	/ɛ/ → /ɒ/	*forget, threat*
	Ae	+	-	/aɪ/ → /eɪ/	*lie*
	B	base and past-participle vowel ≠ past-tense vowel			
	Ba	+	-	base and PP /əʊ/	*grow, know*
	Bb	+	-	base and PP /eɪ/	*shake, take*
	Bc	+	-	base and PP /ɪ/	*forbid, give*
	Bd	+	-		*draw*
	Be	+	-		*fall*
	Bf	+	-		*eat*
	Bg	+	-		*see*
	Bh	+	-		*slay*
	C	all three forms have different vowels			
	Ca	+	-	/aɪ/ → /əʊ/ → /ɪ/	*write, drive*
	Cb	+	-	-	*fly*
	Cc	+	-	-	*do*
	D	+	-	same	*beat*
	E	+	-	/aɪ/ → /aɪ/ or /əʊ/	*dive, thrive*
5		-	+	+	*cut, hit* (all three identical)
6	A	-	+	/iː/ → /ɛ/	*meet, read*
	B	-	+	/ɪ/ → /ʌ/	*dig, spin*
	C	-	+	/aɪ/ → /aʊ/	*find, wind*
	D	-	+	/aɪ/ → /ɪ/	*light, slide*

	E	-	+	/ɪ/ → /æ/	*sit, spit*
	F	-	+	/ɒ/ in past forms	*get, shine*
	G	-	+	/aɪ/ → /ɔː/	*fight*
	H	-	+	/æ/ → /ʊ/	*stand*
	I	-	+	/aɪ/ → /əʊ/	*abide*
	J	-	+	/iː/ → /əʊ/	*heave*
7	A	-	-	/ɪ/ - /æ/ - /ʌ/	*drink, swim*
	B	-	-	base = PP	*come, run*
	C	-	-	-	*go*

Table 42. Quirk et al. (1985)' classification of irregular verbs with all phonological subcategories.

Though rather detailed and restricted, this classification was chosen because it describes phonological changes in past-tense production very similar to Albright and Hayes' (2003) concept of structural analogy. We used this ready-made (and easy to adopt) classification under the assumption that excessive subclassification could make matters confusing for the students, and the seven main categories were retained as the primary division reference with the note on the possible subclassification.

6. Regular verbs were then further divided. Since no previous study has attempted to divide regular verbs into groups based on (structural) analogy, this material builds loosely on Quirk et al.'s classification. Regular verbs are first divided into allomorph groups based on the phonetic quality of the bound morpheme (similarly to Quirk et al.'s verbal identity), then by the number of syllables, and lastly based on the quality of the vowel in the stress syllable (similarly to Quirk et al.'s vowel quality). Each English vowel, diphthong, and triphthong was given a letter code (A to Z – see below), the number of syllables was signified by numbers 1 to 5, and the final allomorph included *t, d*, or *ɪd* coding.

Vowel codes	
e	A
æ	B
ʌ	C
ʊ	D
ɒ	E
ə	F
ɪ	G
i:	H
ɜ:	I
ɔ:	J
u:	K
ɑ:	L
ɪə	M
eə	N
eɪ	O
ɔɪ	P
aɪ	Q
əʊ	R
aʊ	S
ʊə	T
ʊɪ	U
aʊ̯ə̯	V
aɪ̯ə̯	W
ɔɪ̯ə̯	X
eɪə	Y
əʊə	Z

Table 43. Letter codes used for the division of nuclei vowels in the practical teaching application.

Put together, each regular verb received a three-position code that determined its membership in a given analogy group – e.g., the A1 verb *cook* has /ʊ/ as the nuclei vowel in its stressed syllable (which corresponds to letter D), it is a monosyllabic word, and ends in a voiceless *k* (therefore attaching the /t/ allomorph). Its coding will, therefore, be *D1t*. For more examples, see a sample of B2 verbs coded below:

VERB	past tense	past participle	allomorph	vowel	syllables	class
stress	stressed	stressed	t	e	1	A1t
stretch	stretched	stretched	t	e	1	A1t
crack	cracked	cracked	t	æ	1	B1t
dash	dashed	dashed	t	æ	1	B1t
flash	flashed	flashed	t	æ	1	B1t
lack	lacked	lacked	t	æ	1	B1t
sack	sacked	sacked	t	æ	1	B1t
scratch	scratched	scratched	t	æ	1	B1t
slap	slapped	slapped	t	æ	1	B1t
smash	smashed	smashed	t	æ	1	B1t
splash	splashed	splashed	t	æ	1	B1t
stamp	stamped	stamped	t	æ	1	B1t
tap	tapped	tapped	t	æ	1	B1t
blush	blushed	blushed	t	ʌ	1	C1t
bump	bumped	bumped	t	ʌ	1	C1t
punch	punched	punched	t	ʌ	1	C1t
rush	rushed	rushed	t	ʌ	1	C1t
block	blocked	blocked	t	ɒ	1	E1t
chop	chopped	chopped	t	ɒ	1	E1t
shock	shocked	shocked	t	ɒ	1	E1t
squash	squashed	squashed	t	ɒ	1	E1t
blink	blinked	blinked	t	ɪ	1	G1t
dip	dipped	dipped	t	ɪ	1	G1t
grip	gripped	gripped	t	ɪ	1	G1t
itch	itched	itched	t	ɪ	1	G1t
lick	licked	licked	t	ɪ	1	G1t
link	linked	linked	t	ɪ	1	G1t
rip	ripped	ripped	t	ɪ	1	G1t
risk	risked	risked	t	ɪ	1	G1t
skip	skipped	skipped	t	ɪ	1	G1t
switch	switched	switched	t	ɪ	1	G1t
trick	tricked	tricked	t	ɪ	1	G1t
trip	tripped	tripped	t	ɪ	1	G1t

Table 44. An example of the coding used for regular verbs (B2 level, /t/allomorph).

This division of verbs according to language levels, final allomorphs, and structural analogy constitutes the main core of the proposed teaching material (for the full list of regular and irregular verbs for each language level, see Appendix 4). As with Quirk's classification, though, excessive subclassification could sometimes rather hinder the learning process and needs to be handled carefully (the sub-classification might also unnecessarily cause the learner or the instructor to believe that there are actually differences between individual categories within one allomorphic paradigm, other than the quality of the nuclei). The three allomorph groups, therefore, constitute the core groups of the regular verbs, and the verbs should be mainly taught and practiced within their allomorph group rather than primarily in the detailed subclassifications (these can only help the remembrance process). In the future, the regular verbs also need to be more sensitively divided into analogy groups using, for instance, an analogy program based on phonological micro-rules similar to Albright and Hayes' (2003). Further research will also be needed to find supportive evidence for the applicability of the proposed strategy – this could be done by comparing two groups of L2 learners on the acquisition of past-tense forms: One will be taught traditionally, the second one using the analogical verb list; the groups' findings will then be compared to see if there was a significant difference between their performance.

Additional steps can be, of course, taken to make the list even more complex and useful for L2 learners. The verbs can be sorted by frequency (for instance, according to the British National Corpus). Verbs that are phonotactically legal in Czech could be highlighted to distinguish them from verbs that are phonotactically illegal in Czech (and, therefore, potentially more problematic). For this, we would need an updated version of the phonotactics calculator, including syllable nuclei. Since such a counter is not in existence for Czech yet, this division was not used in the thesis. Any specific spelling or pronunciation changes when the regular inflectional morpheme is added can also be highlighted to make the learning process easier for L2 learners.

As Makoe and Shandu (2018: 209) propose, creating a mobile application would be especially suitable for these purposes since such apps are "flexible, accessible, available, and cater for a myriad of interactive activities." Since this is primarily a research thesis and the development of a practical outcome is not its main aim, only the list of verbs will be included as a practical application stemming from the findings of our experiments. However, to raise the effectiveness of the list for L2 teaching, this list should be ideally converted into an online learning application that would focus on the playful, implicit acquisition of past tense, using all the principles for vocabulary learning (such as recycling, spaced repetition, repeated exposure, recontextualization, etc.). Sense support would be specially promoted: Pronunciation would be added to each verb and its past-tense forms, difficult spelling parts will be highlighted, different colours will be used for different analogy groups, etc. Individual verbs will also be studied alone, as well as in a suitable syntactic context to ensure recontextualization of the verb. Exercises could vary from matching the verb with verbs forming the past tense in a similar way, deciding if a word is regular or not, producing the right past-tense form that will fit a sentence, or strengthening the lexical meaning of the verb.

However, since the development of this app will still take a lot of time and financial support, intermediate steps can be taken to get the findings of this study circulated among L2 teachers and used in practice even before the app is ready to be used. The lists could be first made available online for teachers to use, along with guidelines and instructions for the incorporation of the verbs in L2 classrooms. Fun activities will need to be adopted, especially for lower-proficiency students who will not have enough experience with an analogy in English – rhythmic exercises can be used, associations, chants, or drills. Using the list of verbs, more proficient students can create stories (the more unrealistic, the better for memorization) and visualize them. Generally, the use of visualization, rhythmicity, and story-making could be very beneficial for smoother acquisition. Similarly, the app itself will need to be accompanied by further instructions on its proper usage for self-study and classroom learning or for test creation. Our data (especially the qualitative analysis of Experiment 2) have shown the importance of proper instruction when it comes to the pronunciation of the bound morpheme. Attention needs to be paid to the full pronunciation in places where regular pronunciation should be used instead.

Recommendations will also need to be issued for coursebook authors so that the proposed list is included in the supplementary materials in the teacher's books and eventually also in coursebooks. As such, the presented proposal inevitably goes against the traditional approach towards how (ir)regular English past-tense forms are taught: They are usually alphabetical lists of irregular verbs listed at the end of the coursebook. In our proposal, we incline towards analogy, but certain micro-rules are still preserved to help the L2 learner (similarly to Albright & Hayes' structural analogy based on phonological micro-rules). Further research will be, however, needed to find supportive evidence for our hypotheses and to decide whether the proposed division of English verbs is, indeed, beneficial for L2 learners.

6.6 Final reflections

With the first experiment, we have shown that when it comes to morphological perception, L2 learners seem to be sensitive to morphosyntactic information in sublexical items with no access to the lexicon. Our data suggest that morphological decomposition might run parallel to frequency effects, showing that morpheme stripping might operate irrespective of the word's meaning and that certain words might be accessed as units while other words are decomposed on the spot. These phenomena appear to be active from early proficiencies and mirror the results found with native speakers (see Cilibrasi, 2016).

Concerning morphological production, our findings have shown that the L2 production of novel morphologically-inflected forms is characterized by a progressive development from the application of default morphological rules at the lowest language levels to the application of analogy (based on stored examples) at more proficient levels. The language learners, therefore, perform as adult native speakers once they reach a high proficiency level.

In this respect, our findings do not offer a clear answer to the question of how morphological items are stored (as units or as stems and rules) and generated (using single-route or dual-route model). In Experiment 1, we have found evidence both for potential morpheme decomposition and frequency effects; in Experiment 2, we saw a progression from rule application to the use of analogy. Our data, therefore, suggest that both processes are, in each case, active in parallel (see the redundant models discussed at the beginning of this chapter). The choice of one system or the other is governed, for instance, by word frequency.

As far as language transfer is concerned, some L1 effects have been found in all three studies. Experiment 1 has shown that even learners with the lowest proficiencies seemed to be sensitive to English morphology, hinting at the influence of higher morphological awareness related to their L1. Experiment 2 has shown that the lowest language levels (A1 and A2) were influenced by their L1 even in the production of past-tense forms, with items phonotactically legal in Czech promoting significantly quicker reaction times. Experiment 3 has then shown a mild effect of L1 on the whole proficiency scale, once again with quicker reaction times for items phonotactially legal in Czech.

In response to these findings, a practical teaching material has been suggested, and a list of past-tense verbs based on structural analogy has been proposed. Our data and previous research show the potential of promoting the use of analogy from the initial stages of L2 learning, implicitly encouraging the creation of similarity connections and accomplishing an even more effective learning process. Both regular and irregular English verbs were, therefore, divided into similarity groups that could be imported into a mobile learning application. Recommendations for self-study, classroom learning, and coursebook implementation have been suggested. Further research will, however, have to be conducted to prove the effectiveness of these proposals.

These three experiments have enabled us to take a closer look into how English inflectional morphology is handled by L2 learners, both in terms of perception and production. It has also allowed us to look into the development of English inflectional morphology since we recruited participants at different language levels. Given the disproportion in morphological richness between the two languages involved as L1 and L2, these studies have also raised some questions concerning language transfer. The findings of this thesis are also fruitful for teaching practice since they provide specific data about how exactly L2 learners perceive English inflectional morphology, how they work with its rules and how they develop analogical mechanisms, and also to what extent their performance is influenced by their L1. Further research is needed to gain an even more precise understanding; however, the outcomes of this thesis already give us enough material to incorporate some of its results into L2 teaching and into L2 classrooms.

REFERENCES

Agathopoulou, E. (2009). The regular-irregular dissociation in non-native English past tense. *Selected Papers on Theoretical and Applied Linguistics*, *18*, 37–47.

Albright, A. (2002). Islands of Reliability for Regular Morphology: Evidence from Italian. *Language*, *78*(4), 684–709. https://doi.org/10.1353/lan.2003.0002

Albright, A., & Hayes, B. (2003a). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, *90*(2), 119–161. https://doi.org/10.1016/S0010-0277(03)00146-X

Alegre, M., & Gordon, P. (1999). Frequency Effects and the Representational Status of Regular Inflections. *Journal of Memory and Language*, *40*(1), 41–61. https://doi.org/10.1006/jmla.1998.2607

Ambridge, B. (2010). Children's judgments of regular and irregular novel past-tense forms: New data on the English past-tense debate. *Developmental Psychology*, *46*(6), 1497–1504. https://doi.org/10.1037/a0020668

Ambridge, B. (2019). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*. https://doi.org/10.1177/0142723719869731

Anderson, J. D., & Byrd, C. T. (2008). Phonotactic Probability Effects in Children Who Stutter. *Journal of Speech, Language, and Hearing Research*, *51*(4), 851–866. https://doi.org/10.1044/1092-4388(2008/062)

Anisfeld, M., & Gordon, M. (1968). On the psychophonological structure of English inflectional rules. *Journal of Verbal Learning and Verbal Behavior*, *7*(6), 973–979. https://doi.org/10.1016/S0022-5371(68)80055-6

Antolík, T. K. (2020). Ultrasound Tongue Imaging in Second Language Learning. *Studie z Aplikované Lingvistiky*, *11*(1), 109–119.

Avram, L. (2002). *An introduction to language acquisition from a general perspective*. Editura Universitatii Bucuresti.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database*.

Bailey, N., Madden, C., & Krashen, S. D. (1974). Is There A "Natural Sequence" In Adult Second Language Learning? *Language Learning*, *24*(2), 235–243. https://doi.org/10.1111/j.1467-1770.1974.tb00505.x

Basnight-Brown, D. M., Chen, L., Hua, S., Kostić, A., & Feldman, L. B. (2007). Monolingual and bilingual recognition of regular and irregular English verbs: Sensitivity to form similarity varies with first language experience. *Journal of Memory and Language*, *57*(1), 65–80.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using *lme4*. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Beauvillain, C. (1994). Morphological structure in visual word recognition: Evidence from prefixed and suffixed words. *Language and Cognitive Processes*, *9*(3), 317–339.

Beck, M.-L. (1997). Regular verbs, past tense and frequency: Tracking down a potential source of NS/NNS competence differences. *Second Language Research*, 93–115.

Berko, J. (1958). The Child's Learning of English Morphology. *WORD*, *14*(2–3), 150–177. https://doi.org/10.1080/00437956.1958.11659661

Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 489–511. https://doi.org/10.1037//0278-7393.26.2.489

Bittner, D., Dressler, W. U., & Kilani-Schoch, M. (2011). *Development of verb inflection in first language acquisition: A cross-linguistic perspective* (Vol. 21). Walter de Gruyter.

Bliss, H. (2006). L2 acquisition of inflectional morphology: Phonological and morphological transfer effects. *Proceedings of the 8th Generative Approaches to Second Language*

Acquisition Conference (GASLA 2006). Somerville, MA: Cascadilla Proceedings Project, *18*.

Blything, R. P., Ambridge, B., & Lieven, E. V. M. (2018). Children's Acquisition of the English Past-Tense: Evidence for a Single-Route Account From Novel Verb Production Data. *Cognitive Science*. https://doi.org/10.1111/cogs.12581

Bosch, S., Veríssimo, J., & Clahsen, H. (2019). Inflectional morphology in bilingual language processing: An age-of-acquisition study. *Language Acquisition*, *26*(3), 339–360.

Bowerman, M. (1982). Reorganizational processes in lexical and syntactic development. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 319–346). Academic Press.

Brenders, P., Van Hell, J. G., & Dijkstra, T. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of Experimental Child Psychology*, *109*(4), 383–396.

Broe, M. B. (1993). *Specification theory: The treatment of redundancy in generative phonology* [Ph.D., University of Edinburgh]. http://hdl.handle.net/1842/20226

Browman, C. P. (1978). *Tip of the tongue and slip of the ear: Implications for language processing* (Vol. 42). University of California.

Brown, R. (1973). *A First Language The Early Stages*. Harvard University Press.

Brown, R., & McNeill, D. (1966). The 'tip of the tongue' phenomenon. *Journal of Verbal Learning & Verbal Behavior*, *5*(4), 325–337. https://doi.org/10.1016/S0022-5371(66)80040-3

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*(5), 425–455. https://doi.org/10.1080/01690969508407111

Bybee, J. L., & Moder, C. L. (1983). Morphological Classes as Natural Categories. *Language*, *59*(2), 251. https://doi.org/10.2307/413574

Bybee, J. L., & Slobin, D. I. (1982). Rules and Schemas in the Development and Use of the English past Tense. *Language*, *58*(2), 265. https://doi.org/10.2307/414099

Caprin, C., & Guasti, M. T. (2009). The acquisition of morphosyntax in Italian: A cross-sectional study. *Applied Psycholinguistics*, *30*(1), 23–52.

Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society*, *13*(4), 212–214.

Caramazza, A., Laudanna, A., & Romani, C. (1988). Lexical access and inflectional morphology. *Cognition*, *28*(3), 297–332. https://doi.org/10.1016/0010-0277(88)90017-0

Carneiro, M. M. (2017). Investigating bilinguals' sensitivity to English regular past morphology: A self-paced reading experiment with Brazilian learners. *Investigação Da Sensibilidade de Bilíngues à Morfologia Regular de Passado Em Inglês: Um Experimento de Leitura Auto-Cadenciada Com Falantes Brasileiros.*, *17*(3), 483–507. https://doi.org/10.1590/1984-6398201711101

Čechová, M. (1996). *Čeština: Řeč a jazyk*. ISV nakl.

Chen, L., Shu, H. U. A., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject-verb agreement in L2 learning. *Bilingualism*, *10*(2), 161.

Chialant, D., & Caramazza, A. (1995). Where is morphology and how is it processed? The case of written word recognition. In *Morphological aspects of language processing* (pp. 55–76). Lawrence Erlbaum Associates, Inc.

Chomsky, N. (1995). *The Minimalist Program MIT Press Cambridge*. MIT Press.

Cilibrasi, L. (2016). *Word position effects in speech perception*. University of Reading.

Cilibrasi, L., Stojanovik, V., Riddell, P., & Saddy, D. (2019). Sensitivity to Inflectional Morphemes in the Absence of Meaning: Evidence from a Novel Task. *Journal of Psycholinguistic Research*, *48*(3), 747–767. https://doi.org/10.1007/s10936-019-09629-y

Cilibrasi, L., & Tsimpli, I. (2020). Sensitivity to Morphophonological Cues in Monolingual and Bilingual Children: Evidence from a Nonword Task. In *An Anthology of Bilingual Child Phonology*. Multilingual Matters.

Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, *22*(6), 991–1013. https://doi.org/10.1017/S0140525X99002228

Clahsen, H. (2006). Dual-Mechanism Morphology. In *Encyclopedia of Language & Linguistics* (pp. 1–5). Elsevier. https://doi.org/10.1016/B0-08-044854-2/04252-8

Clahsen, H., & Felser, C. (2006). Continuity and shallow structures in language processing: A reply to our commentators. *Applied Psycholinguistics*, *27*(1), 107–126.

Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning*, *60*(1), 21–43.

Clahsen, H., & Neubauer, K. (2010). Morphology, frequency, and the processing of derived words in native and non-native speakers. *Lingua*, *120*(11), 2627–2637.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 1–33). Lawnre ce Erlbaum.

Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, *5*(4), 161–170.

Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, *50*(4), 491–511. https://doi.org/10.1016/j.jml.2004.02.002

Coughlin, C. E., Fiorentino, R., Royle, P., & Steinhauer, K. (2019). Sensitivity to inflectional morphology in a non-native language: Evidence from ERPs. *Frontiers in Communication*, *4*, 21.

Coughlin, C. E., & Tremblay, A. (2015). Morphological decomposition in native and non-native French speakers. *Bilingualism: Language and Cognition*, *18*(3), 524–542.

Crisp, J., & Lambon Ralph, M. A. (2006). Unlocking the nature of the phonological–deep dyslexia continuum: The keys to reading aloud are in phonology and semantics. *Journal of Cognitive Neuroscience*, *18*(3), 348–362.

Cruttenden, A. (1979). *Language in infancy and childhood: A linguistic introduction to language acquisition*. Manchester University Press.

Crystal, D. (2011). *A dictionary of linguistics and phonetics* (Vol. 30). John Wiley & Sons.

Cunnings, I. (2017). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, *20*(4), 659–678.

Cuskley, C., Colaiori, F., Castellano, C., Loreto, V., Pugliese, M., & Tria, F. (2015). The adoption of linguistic rules in native and non-native speakers: Evidence from a Wug task. *Journal of Memory and Language*, *84*, 205–223.

D'Amico, S., Devescovi, A., & Bates, E. (2001). Picture naming and lexical access in Italian children and adults. *Journal of Cognition and Development*, *2*(1), 71–105.

Darcy, I., Yang, M., & Lin, C.-J. C. (2018). *Phonotactics in the bilingual mental lexicon*. https://osf.io/t37x9

de Bot, K., Cox, A., Ralston, S., Schaufeli, A., & Weltens, B. (1995). Lexical processing in bilinguals. *Second Language Research*, *11*(1), 1–19. https://doi.org/10.1177/026765839501100101

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, *55*(S1), 1–25.

Diessel, H. (2013). Construction grammar and first language acquisition. *The Oxford Handbook of Construction Grammar*, *347*, 364.

Dijkstra, T., & Kroll, J. F. (2005). Bilingual visual word recognition and lexical access. *Handbook of Bilingualism: Psycholinguistic Approaches*, *178*, 201.

Dulay, H. C., & Burt, M. K. (1973). Should we teach children syntax? *Language Learning*, *23*(2), 245–258.

Dulay, H. C., & Burt, M. K. (1974). Natural Sequences In Child Second Language Acquisition. *Language Learning*, *24*(1), 37–53. https://doi.org/10.1111/j.1467-1770.1974.tb00234.x

Dušková, L., Malá, M., Šaldová, P., & Brůhová, G. (2006). *Mluvnice současné angličtiny na pozadí češtiny*. Academia.

Ellis, N. C., & Schmidt, R. (1998). Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and Cognitive Processes*, *13*(2–3), 307–336.

Eubank, L., Bischof, J., Huffstutler, A., Leek, P., & West, C. (1997). 'Tom Eats Slowly Cooked Eggs': Thematic-Verb Raising in L2 Knowledge. *Language Acquisition*, *6*(3), 171–199.

Federici, S., Montemagni, S., & Pirrelli, V. (1996). *Analogy-based learning and natural language processing*.

Feldman, L. B., Kostić, A., Basnight-Brown, D. M., Đurđević, D. F., & Pastizzo, M. J. (2010). Morphological facilitation for regular and irregular verb formations in native and non-native speakers: Little evidence for two distinct mechanisms. *Bilingualism (Cambridge, England)*, *13*, 119.

Friedline, B. E. (2011). *Challenges in the second language acquisition of derivational morphology: From theory to practice* [PhD Thesis]. University of Pittsburgh.

Gao, H. (2013). On source language interference in interpretation. *Theory and Practice in Language Studies*, *3*(7), 1194.

Gass, S. M., & Selinker, L. (2013). *Second language acquisition: An introductory course*. Routledge.

Gentner, D., & Namy, L. L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science*, *15*(6), 297–301.

Goad, H., White, L., & Steele, J. (2003). Missing inflection in L2 acquisition: Defective syntax or L1-constrained prosodic representations? *The Canadian Journal of Linguistics/La Revue Canadienne de Linguistique*, *48*(2), 243–263.

Goldschneider, J. M., & DeKeyser, R. M. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning*, *51*(1), 1–50.

Gráf, T. (2015). *Accuracy and Fluency in the Speech of the Advanced Learner of English*. Charles University.

Grainger, J., & Ziegler, J. C. (2011). A Dual-Route Approach to Orthographic Processing. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00054

Guasti, M. T. (1993). Verb syntax in Italian child grammar: Finite and nonfinite verbs. *Language Acquisition*, *3*(1), 1–40.

Guasti, M. T. (2017). *Language acquisition: The growth of grammar* (Second edition). The MIT Press.

Hahne, A., Müller, J., & Clahsen, H. (2003). *Second language learners' processing of inflected words: Behavioral and ERP evidence for storage and decomposition*.

Hakuta, K. (1976). A case study of a Japanese child learning English as a second language 1, 2. *Language Learning*, *26*(2), 321–351.

Halle, M., & Mohanan, K. P. (1985). Segmental phonology of modern English. *Linguistic Inquiry*, *16*(1), 57–116.

Harrington, M. (2006). The lexical decision task as a measure of L2 lexical proficiency. *EUROSLA Yearbook*, *6*(1), 147–168. https://doi.org/10.1075/eurosla.6.10har

Hawkins, R., & Chan, C. Y. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research*, *13*(3), 187–226.

Hawkins, R., & Liszka, S. (2003). Locating the source of defective past tense marking in advanced L2 English speakers. *Language Acquisition and Language Disorders*, *30*, 21–44.

Haznedar, B., & Schwartz, B. D. (1997). Are there optional infinitives in child L2 acquisition. *Proceedings of the 21st Annual Boston University Conference on Language Development*, *21*, 257–268.

Hyams, N. (1992). Morphosyntactic development in Italian and its relevance to parameter-setting models: Comments on the paper by Pizzuto & Caselli. *Journal of Child Language*, *19*(3), 695–709.

Internetová jazyková příručka. (n.d.). Retrieved 29 May 2021, from https://prirucka.ujc.cas.cz/?id=_cite

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language*, *44*(1), 2–31. https://doi.org/10.1006/jmla.2000.2747

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*(4), 603–634. https://doi.org/10.1017/S0142716404001298

Jiang, N. (2007). Selective Integration of Linguistic Knowledge in Adult Second Language Learning. *Language Learning*, *57*(1), 1–33. https://doi.org/10.1111/j.1467-9922.2007.00397.x

Jiang, N. (2012). *Conducting reaction time research in second language studies*. Routledge.

Jiránková, L. (2017). *Word-ending perception in second-language learners of English*. Charles University.

Jiránková, L., Gráf, T., & Kvítková, A. (2019). On the relation between L1 and L2 speech rate. In *Widening the Scope of Learner Corpus Research. Selected Papers from the Fourth Learner Corpus Research Conference* (pp. 19–41). Presses universitaires de Louvain.

Juffs, A. (1998). Some effects of first language argument structure and morphosyntax on second language sentence processing. *Second Language Research*, *14*(4), 406–424.

Karmiloff, K., & Karmiloff-Smith, A. (2001). *Pathways to Language: From Fetus to Adolescent* (1st paperback ed). Harvard Univ. Press.

Kimppa, L., Shtyrov, Y., Hut, S. C. A., Hedlund, L., Leminen, M., & Leminen, A. (2019). Acquisition of L2 morphology by adult language learners. *Cortex*, *116*, 74–90. https://doi.org/10.1016/j.cortex.2019.01.012

Korecky-Kröll, K., Dressler, W. U., Freiberger, E. M., Reinisch, E., Mörth, K., & Libben, G. (2014). Morphonotactic and phonotactic processing in German-speaking adults. *Language Sciences*, *46*, 48–58.

Kroll, J. F. (1993). Accessing conceptual representations for words in a second language. In *The bilingual lexicon* (pp. 53–81). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.6.05kro

Kroll, J. F., & Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. *Practical Aspects of Memory*, *2*(389–395), 8.

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, *18*(2), 137–171. https://doi.org/10.1191/0267658302sr201oa

Ku, Y.-M., & Anderson, R. C. (2003). Development of morphological awareness in Chinese and English. *Reading and Writing*, *16*(5), 399–422. https://doi.org/10.1023/A:1024227231216

Kuczaj, S. A. (1978). Why do children fail to overgeneralize the progressive inflection? *Journal of Child Language*, *5*(1), 167–171. https://doi.org/10.1017/S0305000900002026

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Kwak, S. G., & Kim, J. H. (2017). Central limit theorem: The cornerstone of modern statistics. *Korean Journal of Anesthesiology*, *70*(2), 144–156. https://doi.org/10.4097/kjae.2017.70.2.144

Kwon, E.-Y. (2005). *The "natural order" of morpheme acquisition: A historical survey and discussion of three putative determinants*.

Lardiere, D. (1998a). Case and tense in the 'fossilized' steady state. *Second Language Research*, *14*(1), 1–26.

Lardiere, D. (1998b). Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second Language Research*, *14*(4), 359–375.

Law, F., & Strange, W. (2010). Perception of Canadian French word-final vowels in lexical and morphosyntactic minimal pairs by English learners of French. *The Journal of the Acoustical Society of America*, *128*(4), 2488–2488.

Lehtonen, M., & Laine, M. (2003). How word frequency affects morphological processing in monolinguals and bilinguals. *Bilingualism: Language and Cognition*, *6*(3), 213–225. https://doi.org/10.1017/S1366728903001147

Lehtonen, M., Niska, H., Wande, E., Niemi, J., & Laine, M. (2006). Recognition of Inflected Words in a Morphologically Limited Language: Frequency Effects in Monolinguals and Bilinguals. *Journal of Psycholinguistic Research*, *35*(2), 121. https://doi.org/10.1007/s10936-005-9008-1

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition*, *32*(4), 533–550.

Leonard, L. B., Caselli, M. C., & Devescovi, A. (2002). Italian children's use of verb and noun morphology during the preschool years. *First Language*, *22*(3), 287–304.

Long, M. (1997). Fossilization: Rigor mortis in living linguistic systems. *SLRF97, East Lansing, MI, October*.

Lopatová, L. (2012). *Osvojování cizího jazyka (angličtiny) na základě mateřského jazyka (češtiny) – souvislosti s teorií dětského jazykového obrazu světa*. Charles University.

Luk, Z. P., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural-s, articles, and possessive's. *Language Learning*, *59*(4), 721–754.

Machálek, T. (2020). Kontext: Advanced and flexible corpus query interface. *Proceedings of The 12th Language Resources and Evaluation Conference*, 7003–7008.

Macizo, P., Bajo, T., & Cruz Martín, M. (2010). Inhibitory processes in bilingual language comprehension: Evidence from Spanish–English interlexical homographs. *Journal of Memory and Language*, *63*(2), 232–244. https://doi.org/10.1016/j.jml.2010.04.002

Makoe, M., & Shandu, T. (2018). Developing a Mobile App for Learning English Vocabulary in an Open Distance Learning Context. *The International Review of Research in Open and Distributed Learning*, *19*(4). https://doi.org/10.19173/irrodl.v19i4.3746

Maratsos, M. P. (1987). Gift of tongues: A review of S. Pinker, Language learnability and language development. *London Times Literary Supplement*.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, *57*, i–178.

Marian, V., Blumenfeld, H. K., & Boukrina, O. V. (2008). Sensitivity to phonological similarity within and across languages. *Journal of Psycholinguistic Research*, *37*(3), 141–170.

Marshall, C. R., & van der Lely, H. K. (2006). A challenge to current models of past tense inflection: The impact of phonotactics. *Cognition*, *100*(2), 302–320.

Marshall, C. R., & van der Lely, H. K. J. (2009). Effects of Word Position and Stress on Onset Cluster Production: Evidence from Typical Development, Specific Language Impairment, and Dyslexia. *Language*, *85*(1), 39–57.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 576–585. https://doi.org/10.1037/0096-1523.15.3.576

McQueen, J. M., & Cutler, A. (1998). *Morphology in word recognition*. Blackwell.

Melichar, J., & Styblik, V. (2009). *Český jazyk: Přehled učiva základní školy s cvičeními a klíčem*. Fortuna.

Moore, D. S., & McCabe, G. P. (1999). *Introduction to the practice of statistics* (3rd ed). W.H. Freeman.

Murphy, V. A. (2000). *Inflectional morphology and second language learning systems: An investigation of the dual-mechanism model and L2 morphology*. McGili University.

Nooteboom, S. G. (1981). Lexical retrieval from fragments of spoken words: Beginnings vs endings. *Journal of Phonetics*, *9*(4), 407–424. https://doi.org/10.1016/S0095-4470(19)31017-4

Nosofsky, R. M. (1990). Relations between exemplar-similarity and likelihood models of classification. *Journal of Mathematical Psychology*, *34*(4), 393–418. https://doi.org/10.1016/0022-2496(90)90020-A

Oshodi, B. (2014). Assessing the so called marked inflectional features of Nigerian English: A second language acquisition theory account. *Íkala, Revista de Lenguaje y Cultura*, *19*(1), 15–26.

Osolsobě, K. (2017). *Flexe*. CzechEncy - Nový Encyklopedický Slovník Češtiny. https://www.czechency.org/slovnik/FLEXE

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Peretokina, V., Best, C. T., Tyler, M. D., Shaw, J. A., & Di Biase, B. (2015). Perception of English codas in various phonological and morphological contexts by Mandarin learners of English. *ICPhS*.

Peretokina, V., Tyler, M. D., & Best, C. T. (2014). Influence of phonological, morphological, and prosodic factors on phoneme detection by native and second-language adults. *Proceedings of the 15th Australasian International Conference on Speech Science and Technology (SST2014), 2-5 December 2014, Rydges Latimer Hotel, Christchurch, New Zealand*, 211–214.

Pérez, A., Joseph, H. S. S. L., Bajo, T., & Nation, K. (2016). Evaluation and revision of inferential comprehension in narrative texts: An eye movement study. *Language, Cognition and Neuroscience*, *31*(4), 549–566. https://doi.org/10.1080/23273798.2015.1115883

Pinker, S. (1991). Rules of language. *Science*, *253*(5019), 530–535. https://doi.org/10.1126/science.1857983

Pinker, S., & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences*, *6*(11), 456–463. https://doi.org/10.1016/S1364-6613(02)01990-3

Pitt, M. A. (1998). Phonological processes and the perception of phonotactically illegal consonant clusters. *Perception & Psychophysics*, *60*(6), 941–951. https://doi.org/10.3758/BF03211930

Pitt, M. A., & Samuel, A. G. (1995). Lexical and Sublexical Feedback in Auditory Word Recognition. *Cognitive Psychology*, *29*(2), 149–188. https://doi.org/10.1006/cogp.1995.1014

Pizzuto, E., & Caselli, M. C. (1992). The acquisition of Italian morphology: Implications for models of language development. *Journal of Child Language*, *19*(3), 491–557.

Plag, I. (2003). *Word-formation in English*. Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511841323

Poeppel, D., & Wexler, K. (1993). The full competence hypothesis of clause structure in early German. *Language*, 1–33.

Portin, M., Lehtonen, M., & Laine, M. (2006). Processing of inflected nouns in late bilinguals. *Applied Psycholinguistics*, *28*, 135–156. https://doi.org/10.1017/S014271640607007X

Post, B., Marslen-Wilson, W. D., Randall, B., & Tyler, L. K. (2008). The processing of English regular inflections: Phonological cues to morphological structure. *Cognition*, *109*(1), 1–17. https://doi.org/10.1016/j.cognition.2008.06.011

Potter, M. C., So, K.-F., Eckardt, B. V., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, *23*(1), 23–38. https://doi.org/10.1016/S0022-5371(84)90489-4

Prasada, S., & Pinker, S. (1993). Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes*, *8*(1), 1–56. https://doi.org/10.1080/01690969308406948

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, *16*(2), 103–133.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (Eds.). (1985). *A Comprehensive grammar of the English language*. Longman.

R Development Core Team. (2011). *R: A language and environment for statistical computing: Reference index*. R Foundation for Statistical Computing. http://www.polsci.wvu.edu/duval/PS603/Notes/R/fullrefman.pdf

Ramscar, M. (2002). The role of meaning in inflection: Why the past tense does not require a rule. *Cognitive Psychology*, *45*(1), 45–94. https://doi.org/10.1016/S0010-0285(02)00001-4

Roach, P. (2009). *English phonetics and phonology: A practical course* (4th ed). Cambridge University Press.

Rosenbaum, N. (2011). Sunday Morning Exercise: Take 'The Wug Test.'. In *On Being*.

Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 2: Psychological and biological models* (pp. 216–271). MIT Press.

Sandra, D. (1994). The morphology of the mental lexicon: Internal word structure viewed from a psycholinguistic perspective. *Language and Cognitive Processes*, *9*(3), 227–269. https://doi.org/10.1080/01690969408402119

Sasisekaran, J., Smith, A., Sadagopan, N., & Weber-Fox, C. (2010). Nonword Repetition in Children and Adults: Effects on Movement Coordination. *Developmental Science*, *13*, 521–532. https://doi.org/10.1111/j.1467-7687.2009.00911.x

Sato, M. (2007). *Sensitivity to syntactic and semantic information in second language sentence processing* [PhD Thesis]. University of Essex.

Schoonbaert, S., Duyck, W., Brysbaert, M., & Hartsuiker, R. J. (2009). Semantic and translation priming from a first language to a second and back: Making sense of the findings. *Memory & Cognition*, *37*(5), 569–586. https://doi.org/10.3758/MC.37.5.569

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, *10*, 209–231.

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, *11*(02).

Slobin, D. I. (1971). Developmental Psycholinguistics. In *A survey of linguistic science*. University of Maryland Press.

Šmilauer, V. (1973). *Nauka o českém jazyku*. SPN.

Smolík, F. (2002). Osvojování českých slovesných tvarů v raném věku. *Československá Psychologie*, *46*(5), 450–461.

Solt, S., Pugach, Y., Klein, E., Adams, K., Stoyneshka, I., & Rose, T. (2003). L2 Perception and Production of the English Regular Past: Evidence of Phonological Effects. *Proceedings of the Annual Boston University Conference on Language Development*, *28*.

Stemberger, J. P. (1981). Morphological Haplology. *Language*, *57*(4), 791. https://doi.org/10.2307/414242

Stemberger, J. P., & MacWhinney, B. (1986). Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition*, *14*(1), 17–26. https://doi.org/10.3758/BF03209225

Šturm, P., & Lukeš, D. (2017). Fonotaktická analýza obsahu slabik na okrajích českých slov v mluvené a psané řeči. *Slovo a Slovesnost*, *78*(2), 99–118.

Sunderman, G., & Kroll, J. F. (2006). First Language Activation During Second Language Lexical Processing: An Investigation of Lexical Form, Meaning, and Grammatical Class. *Studies in Second Language Acquisition*, *28*(3), 387–422. https://doi.org/10.1017/S0272263106060177

Talamas, A., Kroll, J. F., & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second-language vocabulary. *Bilingualism: Language and Cognition*, *2*(1), 45–58.

Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, *11*(1–2). https://doi.org/10.1515/cogl.2001.012

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, *30*(1), 37–69.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. *Mind and Context in Adult Second Language Acquisition*, *2005*, 141–178.

Van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*(3), 458–483.

Vitevitch, M. S., & Luce, P. A. (2004). A Web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 481–487. https://doi.org/10.3758/BF03195594

Voga, M., Anastassiadis-Syméonidis, A., & Giraudo, H. (2014). Does morphology play a role in L2 processing?: Two masked priming experiments with Greek speakers of ESL. *Lingvisticæ Investigationes*, *37*(2), 338–352. https://doi.org/10.1075/li.37.2.10vog

Westermann, G., & Ruh, N. (2012). A neuroconstructivist model of past tense development and processing. *Psychological Review*, *119*(3), 649–667. https://doi.org/10.1037/a0028258

White, L. (2003). Fossilization in steady state L2 grammars: Persistent problems with inflectional morphology. *Bilingualism*, *6*(2), 129.

Xu, F., & Pinker, S. (1995). Weird past tense forms. *Journal of Child Language*, *22*(03). https://doi.org/10.1017/S0305000900009946

Yoke, S. K., & Hasan, N. H. (2014). Analogy as a Tool for the Acquisition of English Verb Tenses among Low Proficiency L2 Learners. *English Language Teaching*, *7*(4), p46. https://doi.org/10.5539/elt.v7n4p46

APPENDIX

Appendix 1 – Experiment 1: The perception experiment

Table 45. Novel words used in Experiment 1 (adopted from Cilibrasi, 2016).

Stem number	Morpho	Morpho	Non-morpho	Non-morpho	Control	Control
1	1 vɪld	2 vɪlz	41 vɪlt	42 vɪls	81 vɪlb	82 vɪlm
2	3 vɛld	4 vɛlz	43 vɛlt	44 vɛls	83 vɛlb	84 vɛlm
3	5 væld	6 vælz	45 vælt	46 væls	85 vælb	86 vælm
4	7 vɔld	8 vɔlz	47 vɔlt	48 vɔls	87 vɔlb	88 vɔlm
5	9 vʌld	10 vʌlz	49 vʌlt	50 vʌls	89 vʌlb	90 vʌlm
6	11 nɪld	12 nɪlz	51 nɪlt	52 nɪls	91 nɪlb	92 nɪlm
7	13 naɪld	14 naɪlz	53 naɪlt	54 naɪls	93 naɪlb	94 naɪlm
8	15 næld	16 nælz	55 nælt	56 næls	95 nælb	96 nælm
9	17 nɔld	18 nɔlz	57 nɔlt	58 nɔls	97 nɔlb	98 nɔlm
10	19 nʌld	20 nʌlz	59 nʌlt	60 nʌls	99 nʌlb	100 nʌlm
11	21 θɪld	22 θɪlz	61 θɪlt	62 θɪls	101 θɪlb	102 θɪlm
12	23 θaɪld	24 θaɪlz	63 θaɪlt	64 θaɪls	103 θaɪlb	104 θaɪlm
13	25 θæld	26 θælz	65 θælt	66 θæls	105 θælb	106 θælm
14	27 θɔld	28 θɔlz	67 θɔlt	68 θɔls	107 θɔlb	108 θɔlm
15	29 θʌld	30 θʌlz	69 θʌlt	70 θʌls	109 θʌlb	110 θʌlm
16	31 dʒɑld	32 dʒɑlz	71 dʒɑlt	72 dʒɑls	111 dʒɑlb	112 dʒɑlm
17	33 dʒaɪld	34 dʒaɪlz	73 dʒaɪlt	74 dʒaɪls	113 dʒaɪlb	114 dʒaɪlm
18	35 dʒæld	36 dʒælz	75 dʒælt	76 dʒæls	115 dʒælb	116 dʒælm
19	37 dʒɔld	38 dʒɔlz	77 dʒɔlt	78 dʒɔls	117 dʒɔlb	118 dʒɔlm
20	39 dʒʌld	40 dʒʌlz	79 dʒʌlt	80 dʒʌls	119 dʒʌlb	120 dʒʌlm

Table 46. PSF and BSF values used in Experiment 1 (adopted from Cilibrasi, 2016).

Klattese transcription	Positional segment frequency (PSF)	Biphone segment frequency (BSF)
v@lz	1,2044	1,013
vclz	1,1811	1,0124
vIlz	1,1876	1,0113
vElz	1,1247	1,0039
v@lz	1,2044	1,013
vclz	1,1811	1,0124
nclz	1,1876	1,0113

n^lz	1,1247	1,0039
TIlz	1,1261	1,004
TYlz	1,1488	1,0062
JYlz	1,1887	1,0102
J@lz	1,1268	1,0031
Jclz	1,1338	1,0032
J^lz	1,1789	1,0099
T@lz	1,116	1,0039
Tclz	1,1387	1,0062
T^lz	1,172	1,009
Jalz	1,109	1,0039
vIls	1,1317	1,0052
vEls	1,16	1,0072
v@ls	1,2424	1,0146
vcls	1,219	1,014
vIls	1,2256	1,0129
vEls	1,1627	1,0054
v@ls	1,2424	1,0146
vcls	1,219	1,014
ncls	1,2256	1,0129
n^ls	1,1627	1,0054
TIls	1,1641	1,0055
Tyls	1,1868	1,0077
JYls	1,2267	1,0118
J@ls	1,1648	1,0046
Jcls	1,1718	1,0048
J^ls	1,2169	1,0115
T@ls	1,154	1,0054
Tcls	1,1767	1,0077
T^ls	1,21	1,0106
Jals	1,147	1,0055
vIlm	1,1697	1,0067
vElm	1,198	1,0088
v@lm	1,2218	1,0139
vclm	1,1985	1,0132
vIlm	1,205	1,0121
vElm	1,1421	1,0047
v@lm	1,2218	1,0139
vclm	1,1985	1,0132
nclm	1,205	1,0121

n^lm	1,1421	1,0047
TIlm	1,1435	1,0048
TYlm	1,1662	1,007
T@lm	1,2062	1,011
Tclm	1,1442	1,0039
T^lm	1,1894	1,0098
Jalm	1,1265	1,0047
JYlm	1,1492	1,006
J@lm	1,1775	1,008
Jclm	1,1512	1,004
J^lm	1,1964	1,0107
v@lz	1,1334	1,0047
vclz	1,1562	1,007

Table 47. The overview of the linear mixed-effects model run for the decompositional analysis. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Fixed effects	**Estimate**	**SE**	**t-value**	**Pr(>|t|)**
Intercept	8.095e-01	8.719e-02	9.284	< .001 ***
cond2 (non-morpho)	1.648e-01	8.433e-02	1.954	0.053 .
cond3 (control)	2.517e-01	8.398e-02	2.997	0.00328 **
types (identical items)	2.013e-02	7.819e-03	2.574	0.01133 *
length	7.426e-01	8.939e-02	8.308	< .001 ***
levelA1	9.589e-02	6.036e-02	1.589	0.115
levelA2	4.979e-02	6.076e-02	0.820	0.414
levelB1	-2.227e-02	5.915e-02	-0.377	0.707
levelB2	-5.896e-02	5.914e-02	-0.997	0.321
levelC1	-6.123e-02	5.914e-02	-1.035	0.303
cond2:length	-3.898e-01	1.165e-01	-3.346	0.001 **
cond3:length	-3.555e-01	1.127e-01	-3.155	0.002 **
cond2:levelA1	-1.953e-02	2.112e-02	-0.925	0.355
cond3:levelA1	-1.993e-03	2.182e-02	-0.091	0.927
cond2:levelA2	-1.670e-02	2.123e-02	-0.787	0. 432
cond3:levelA2	-8.311e-03	2.194e-02	-0.379	0.705
cond2:levelB1	-1.114e-02	2.070e-02	-0.538	0.590

cond3:levelB1	-2.457e-02	2.139e-02	-1.149	0.251
cond2:levelB2	-2.407e-02	2.066e-02	-1.165	0.244
cond3:levelB2	-4.154e-02	2.133e-02	-1.948	0.051 .
cond2:levelC1	-6.482e-03	2.067e-02	-0.314	0.754
cond3:levelC1	-2.515e-02	2.135e-02	-1.178	0.239

Table 48. The overview of the linear mixed-effects model run for the atomist analysis with PSF values. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Fixed effects	**Estimate**	**SE**	**t-value**	**Pr(>|t|)**
Intercept	2.258e+00	3.810e-01	5.927	< .001 ***
levelA1	1.120e-01	3.097e-01	0.362	0.7176
levelA2	1.559e-01	3.113e-01	0.501	0.6167
levelB1	-9.546e-02	3.038e-01	-0.314	0.7533
levelB2	-5.253e-02	3.028e-01	-0.173	0. 8623
levelC1	-1.206e-01	3.030e-01	-0.398	0.6906
gPSF	-8.087e-01	3.210e-01	-2.519	0.0121 *
levelA1:PSF	-1.977e-02	2.586e-01	-0.076	0.9391
levelA2:PSF	-9.727e-02	2.600e-01	-0.374	0.7083
levelB1:PSF	5.242e-02	2.536e-01	0.207	0.8363
levelB2:PSF	-2.389e-02	2.528e-01	-0.095	0.9247
levelC1:PSF	4.189e-02	2.530e-01	0.166	0.8685

Table 49. The overview of the linear mixed-effects model run for the atomist analysis with BF values. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Fixed effects	**Estimate**	**SE**	**t-value**	**Pr(>|t|)**
Intercept	3.005e+00	3.072e+00	0.978	0.329
levelA1	7.981e-02	2.407e+00	0.033	0.974
levelA2	7.204e-01	2.421e+00	0.298	0.766
levelB1	2.511e-01	-2.361e+00	-0.106	0.915
levelB2	-1.828e-01	2.354e+00	-0.078	0.938
levelC1	-3.383e-01	2.356e+00	-0.144	0.886

BF	-1.682e+00	3.046e+00	-0.552	0.581
levelA1:BF	8.876e-03	2.386e+00	0.004	0.997
levelA2:BF	-6.732e-01	2.400e+00	-0.280	0.779
levelB1:BF	2.154e-01	2.340e+00	0.092	0.927
levelB2:BF	1.014e-01	2.333e+00	0.043	0.965
levelC1:BF	2.646e-01	2.336e+00	0.113	0.910

Table 50. The overview of the generalized linear mixed-effects model run for the accuracy analysis. Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1.

Fixed effects	**Estimate**	**SE**	**z-value**	**Pr(>|z|)**
Intercept	1.940	0.316	6.144	< .001 ***
cond2 (non-morpho)	0.159	0.245	0.650	0.516
cond3 (control)	-1.139	0.208	-5.481	< .001 ***
typesame (identical items)	-5.737	0.493	-11.627	< .001 ***
levelA1	1.889	0.347	5.447	< .001 ***
levelA2	2.091	0.360	5.813	< .001 ***
levelB1	2.003	0.337	5.939	< .001 ***
levelB2	2.131	0.341	6.255	< .001 ***
levelC1	2.094	0.340	6.155	< .001 ***
cond2:typesame	0.119	0.345	0.345	0.730
cond3:typesame	1.318	0.323	4.082	< .001 ***
typesame:levelA1	-2.058	0.519	-3.963	< .001 ***
typesame:levelA2	-2.173	0.527	-4.120	< .001 ***
typesame:levelB1	-1.999	0.501	-3.989	< .001 ***
typesame:levelB2	-2.304	0.508	-4.539	< .001 ***
typesame:levelC1	-2.631	0.522	-5.038	< .001 ***

Appendix 2 – Experiment 2: The production experiment

Table 51. Novel verbs used in Experiment 2 (adopted from Blything et al., 2018).

No.	Verb	Pronunciation	Set	Similarity to regulars	Similarity to irregulars
1	blafe	bleɪf	A	0.998	0.000
2	dape	deɪp	A	0.993	0.000
3	nace	neɪs	A	0.998	0.046
4	tesh	teʃ	A	0.998	0.000
5	chake	t͡ʃeɪk	A	0.900	0.835
6	drit	drɪt	A	0.944	0.477
7	plim	plɪm	A	0.872	0.348
8	teep	ti:p	A	0.963	0.559
9	drice	draɪs	A	0.998	0.848
10	glip	glɪp	A	0.988	0.059
11	stin	stɪn	A	0.972	0.280
12	stip	stɪp	A	0.988	0.149
13	pank	pæŋk	A	0.958	0.000
14	preak	pri:k	A	0.941	0.033
15	rask	ra:sk	A	0.982	0.000
16	trisk	trɪsk	A	0.963	0.559
17	bredge	bred͡ʒ	B	0.995	0.000
18	gezz	gez	B	0.988	0.000
19	stire	staɪə	B	0.985	0.000
20	wiss	wɪs	B	0.998	0.000
21	blig	blɪg	B	0.961	0.880
22	gleed	gli:d	B	0.872	0.000
23	skride	skraɪd	B	0.887	0.731
24	spling	splɪŋ	B	0.925	0.880
25	bize	baɪz	B	0.988	0.121
26	dize	daɪz	B	0.988	0.554
27	flidge	flɪd͡ʒ	B	0.995	0.200
28	gare	geə	B	0.985	0.257
29	gude	gu:d	B	0.989	0.014
30	nold	nəʊld	B	0.900	0.014
31	nung	nʌŋ	B	0.925	0.000
32	shilk	ʃɪlk	B	0.982	0.000

Table 52. Frame sentences used in Experiment 2 (adopted from Blything et al., 2018).

1. Her friend likes to bize.
   Look, there he is bizing.
   Everyday he bizes.
   So yesterday he. . .

2. My dad likes to bize.
   Look, there he is bizing.
   Everyday he bizes.
   So yesterday he. . .

3. The animal likes to bize.
   Look, there it is bizing.
   Everyday it bizes.
   So yesterday it. . .

4. My dad likes to blafe.
   Look, there he is blafing.
   Everyday he blafes.
   So yesterday he. . .

5. The animal likes to blafe.
   Look, there it is blafing.
   Everyday it blafes.
   So yesterday it. . .

6. The waitress likes to blafe.
   Look, there she is blafing.
   Everyday she blafes.
   So yesterday she. . .

7. The animal likes to blig.
   Look, there it is bligging.
   Everyday it bligs.
   So yesterday it. . .

8. The waitress likes to blig.
   Look, there she is bligging.
   Everyday she bligs.
   So yesterday she. . .

9. The dog likes to blig.
   Look, there it is bligging.
   Everyday it bligs.
   So yesterday it. . .

10. The waitress likes to bredge.
    Look, there she is bredging.
    Everyday she bredges.
    So yesterday she. . .

11. The dog likes to bredge.
    Look, there it is bredging.
    Everyday it bredges.
    So yesterday it. . .

12. The group likes to bredge.
    Look, there it is bredging.
    Everyday it bredges.
    So yesterday it. . .

13. The dog likes to chake.
    Look, there it is chaking.
    Everyday it chakes.
    So yesterday it. . .

14. The group likes to chake.
    Look, there it is chaking.
    Everyday it chakes.
    So yesterday it. . .

15. The baby likes to chake.
    Look, there she is chaking.
    Everyday she chakes.
    So yesterday she. . .

16. The group likes to dape.
    Look, there it is daping.
    Everyday it dapes.
    So yesterday it. . .

17. The baby likes to dape.
    Look, there she is daping.
    Everyday she dapes.
    So yesterday she. . .

18. My father likes to dape.
    Look, there he is daping.
    Everyday he dapes.
    So yesterday he. . .

19. The baby likes to dize.
Look, there she is dizing.
Everyday she dizes.
So yesterday she. . .

20. My father likes to dize.
Look, there he is dizing.
Everyday he dizes.
So yesterday he. . .

21. The parent likes to dize.
Look, there he is dizing.
Everyday he dizes.
So yesterday he. . .

22. My father likes to drice.
Look, there he is dizing.
Everyday he dizes.
So yesterday he. . .

23. The parent likes to drice.
Look, there he is dizing.
Everyday he dizes.
So yesterday he. . .

24. My mum likes to drice.
Look, there she is dricing.
Everyday she drices.
So yesterday she. . .

25. The parent likes to drit.
Look, there he is dritting.
Everyday he drits.
So yesterday he. . .

26. My mum likes to drit.
Look, there she is dritting.
Everyday she drits.
So yesterday she. . .

27. The taxi driver likes to drit.
Look, there he is dritting.
Everydayshe drits.
So yesterday he. . .

28. My mum likes to fligdge.
Look, there she is fligding.
Everyday she fligdes.
So yesterday she. . .

29. The taxi driver likes to fligde.
Look, there he is fligding.
Everyday he fligdes.
So yesterday he. . .

30. The doctor likes to fligde.
Look, there he is fligding.
Everyday he fligdes.
So yesterday he. . .

31. The taxi driver likes to gare.
Look, there he is garing.
Everyday he gares.
So yesterday he. . .

32. The doctor likes to gare.
Look, there he is garing.
Everyday he gares.
So yesterday he. . .

33. The family likes to gare.
Look, there it is garing.
Everyday it gares.
So yesterday it. . .

34. The doctor likes to gezz.
Look, there he is gezzing.
Everyday he gezzes.
So yesterday he. . .

35. The family likes to gezz.
Look, there it is gezzing.
Everyday it gezzes.
So yesterday it. . .

36. My mother likes to gezz.
Look, there she is gezzing.
Everyday she gezzes.
So yesterday she. . .

37. The family likes to gleed.
Look, there it is gleeding.
Everyday it gleeds.
So yesterday it. . .

38. My mother likes to gleed.
Look, there she is gleeding.
Everyday she gleeds.
So yesterday she. . .

39. His wife likes to gleed.
Look, there she is gleeding.
Everyday she gleeds.
So yesterday she. . .

40. My mother likes to glip.
Look, there she is glipping.
Everyday she glips.
So yesterday she. . .

41. His wife likes to glip.
Look, there she is glipping.
Everyday she glips.
So yesterday she. . .

42. His daughter likes to glip.
Look, there she is glipping.
Everyday she glips.
So yesterday she. . .

43. His wife likes to gude.
Look, there she is guding.
Everyday she gudes.
So yesterday she. . .

44. His daughter likes to gude.
Look, there she is guding.
Everyday she gudes.
So yesterday she. . .

45. My pet likes to gude.
Look, there it is guding.
Everyday it gudes.
So yesterday it. . .

46. His daughter likes to nace.
Look, there she is nacing.
Everyday she naces.
So yesterday she. . .

47. My pet likes to nace.
Look, there it is nacing.
Everyday it naces.
So yesterday it. . .

48. The girl likes to nace.
Look, there she is nacing.
Everyday she naces.
So yesterday she. . .

49. My pet likes to nold.
Look, there it is nolding.
Everyday it nolds.
So yesterday it. . .

50. The girl likes to nold.
Look, there she is nolding.
Everyday she nolds.
So yesterday she. . .

51. The woman likes to nold.
Look, there she is nolding.
Everyday she nolds.
So yesterday she. . .

52. The girl likes to nung.
Look, there she is nunging.
Everyday she nungs.
So yesterday she. . .

53. The woman likes to nung.
Look, there she is nunging.
Everyday she nungs.
So yesterday she. . .

54. The band likes to nung.
Look, there it is nunging.
Everyday it nungs.
So yesterday it. . .

55. The woman likes to pank.
Look, there she is panking.
Everyday she panks.
So yesterday she. . .

56. The band likes to pank.
Look, there it is panking.
Everyday it panks.
So yesterday it. . .

57. Her son likes to pank.
Look, there he is panking.
Everyday he panks.
So yesterday he. . .

58. The band likes to plim.
Look, there it is plimming.
Everyday it plims.
So yesterday it. . .

59. Her son likes to plim.
Look, there he is plimming.
Everyday he plims.
So yesterday he. . .

60. The boy likes to plim.
Look, there he is plimming.
Everyday he plims.
So yesterday he. . .

61. Her son likes to preak.
Look, there he is preaking.
Everyday he preaks.
So yesterday he. . .

62. The boy likes to preak.
Look, there he is preaking.
Everyday he preaks.
So yesterday he. . .

63. The man likes to preak.
Look, there he is preaking.
Everyday he preaks.
So yesterday he. . .

64. The boy likes to rask.
Look, there he is rasking.
Everyday he rasks.
So yesterday he. . .

65. The man likes to rask.
Look, there he is rasking.
Everyday he rasks.
So yesterday he. . .

66. The student likes to rask.
Look, there he is rasking.
Everyday he rasks.
So yesterday he. . .

67. The man likes to shilk.
Look, there he is shilking.
Everyday he shilks.
So yesterday he. . .

68. The student likes to shilk.
Look, there he is shilking.
Everyday he shilks.
So yesterday he. . .

69. My sister likes to shilk.
Look, there she is shilking.
Everyday she shilks.
So yesterday she. . .

70. The student likes to skride.
Look, there he is skriding.
Everyday he skrides.
So yesterday he. . .

71. My sister likes to skride.
Look, there she is skriding.
Everyday she skrides.
So yesterday she. . .

72. The child likes to skride.
Look, there he is skriding.
Everyday he skrides.
So yesterday he. . .

73. My sister likes to spling.
Look, there she is splinging.
Everyday she splings.
So yesterday she. . .

74. The child likes to spling.
Look, there he is splinging.
Everyday he splings.
So yesterday he. . .

75. The waiter likes to spling.
Look, there he is splinging.
Everyday he splings.
So yesterday he. . .

76. The child likes to stin.
Look, there he is stinning.
Everyday he stins.
So yesterday he. . .

77. The waiter likes to stin.
Look, there he is stinning.
Everyday he stins.
So yesterday he. . .

78. The teacher likes to stin.
Look, there he is stinning.
Everyday he stins.
So yesterday he. . .

79. The waiter likes to stip.
Look, there he is stipping.
Everyday he stips.
So yesterday he. . .

80. The teacher likes to stip.
Look, there he is stipping.
Everyday he stips.
So yesterday he. . .

81. This person likes to stip.
Look, there she is stipping.
Everyday she stips.
So yesterday she. . .

82. The teacher likes to stire.
Look, there he is stiring.
Everyday he stires.
So yesterday he. . .

83. This person likes to stire.
Look, there she is stiring.
Everyday she stires.
So yesterday she. . .

84. The cat likes to stire.
Look, there it is stiring.
Everyday it stires.
So yesterday it. . .

85. This person likes to teep.
Look, there she is teepning.
Everyday she teeps.
So yesterday she. . .

86. The cat likes to teep.
Look, there it is teepning.
Everyday it teeps.
So yesterday it. . .

87. Jane’s husband likes to teep.
Look, there he is teepning.
Everyday he teeps.
So yesterday he. . .

88. The cat likes to tesh.
Look, there it is teshing.
Everyday it teshes.
So yesterday it. . .

89. Jane’s husband likes to tesh.
Look, there he is teshing.
Everyday he teshes.
So yesterday he. . .

90. My brother likes to tesh.
Look, there he is teshing.
Everyday he teshes.
So yesterday he. . .

91. Jane’s husband likes to trisk.
Look, there he is trisking.
Everyday he trisks.
So yesterday he. . .

92. My brother likes to trisk.
Look, there he is trisking.
Everyday he trisks.
So yesterday he. . .

93. Her friend likes to trisk.
Look, there he is trisking.
Everyday he trisks.
So yesterday he. . .

94. My brother likes to wiss.
Look, there he is wissing.
Everyday he wisses.
So yesterday he. . .

95. Her friend likes to wiss.
Look, there he is wissing.
Everyday he wisses.
So yesterday he. . .

96. My dad likes to wiss.
Look, there he is wissing.
Everyday he wisses.
So yesterday he. . .

Table 53. The overview of the generalized linear mixed-effects model run for the initial analysis of produced forms. Significance codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1.

Fixed effects	**Estimate**	**SE**	**z-value**	**Pr(>|z|)**
Intercept	-0.736	7.679	-0.096	0.923
LevelA2	-1.215	7.811	-0.156	0.876
LevelB1	-2.647	6.959	-0.380	0.703
LevelB2	-8.713	6.921	-1.259	0.208

LevelC1	-5.945	6.711	-0.886	0.375
LevelNS (native speakers)	-3.174	7.454	-0.426	0.670
MGL_R (similarity to regulars)	4.139	7.903	0.524	0.600
MGL_IR (similarity to irregulars)	21.362	23.309	0.916	0.359
LevelA2:MGL_R	0.974	8.043	0.121	0.903
LevelB1:MGL_R	1.857	7.158	0.259	0.795
LevelB2:MGL_R	8.178	7.131	1.147	0.251
LevelC1:MGL_R	5.154	6.907	0.746	0.455
LevelNS:MGL_R	1.131	7.668	0.147	0.882
LevelA2:MGL_IR	-13.339	24.933	-0.535	0.592
LevelB1:MGL_IR	-16.970	22.897	-0.741	0.458
LevelB2:MGL_IR	-8.091	23.002	-0.352	0.725
LevelC1:MGL_IR	-1.992	21.523	-0.093	0.926
LevelNS:MGL_IR	-51.684	23.612	-2.189	0.028 *
MGL_R:MGL_IR	-20.711	24.283	-0.853	0.393
LevelA2:MGL_R:MGL_IR	12.186	26.004	0.469	0.639
LevelB1:MGL_R:MGL_IR	16.508	23.863	0.692	0.489
LevelB2:MGL_R:MGL_IR	6.129	23.963	0.256	0.798
LevelC1:MGL_R:MGL_IR	-0.674	22.394	-0.030	0.976
LevelNS:MGL_R:MGL_IR	51.647	24.626	2.097	0.036 *

Table 54. The overview of the generalized linear mixed-effects model run for the A1 post-hoc analysis.

Fixed effects	**Estimate**	**SE**	**z-value**	**Pr(>|z|)**
Intercept	4.731	-13.089	-0.361	0.718
MGL_R (similarity to regulars)	8.830	13.584	0.650	0.516
MGL_IR (similarity to irregulars)	43.081	47.149	0.914	0.361
MGL_R:MGL_IR	-42.433	48.707	-0.871	0.384

Table 55. The overview of the generalized linear mixed-effects model run for the A2 post-hoc analysis.

Fixed effects	Estimate	SE	z-value	Pr(>|z|)
Intercept	-2.420	9.035	-0.268	0.789
MGL_R (similarity to regulars)	5.697	9.348	0.609	0.542
MGL_IR (similarity to irregulars)	6.469	22.683	0.285	0.775
MGL_R:MGL_IR	-6.752	23.844	-0.283	0.777

Table 56. The overview of the generalized linear mixed-effects model run for the B1 post-hoc analysis.

Fixed effects	Estimate	SE	z-value	Pr(>|z|)
Intercept	-5.509	8.346	-0.660	0.509
MGL_R (similarity to regulars)	8.339	8.627	0.967	0.334
MGL_IR (similarity to irregulars)	7.535	21.718	0.347	0.729
MGL_R:MGL_IR	-7.323	22.819	-0.321	0.748

Table 57. The overview of the generalized linear mixed-effects model run for the B2 post-hoc analysis.

Fixed effects	Estimate	SE	z-value	Pr(>|z|)
Intercept	-8.701	5.543	-1.570	0.117
MGL_R (similarity to regulars)	11.509	5.731	2.008	0.045 *
MGL_IR (similarity to irregulars)	12.130	13.218	0.918	0.359
MGL_R:MGL_IR	-13.360	13.891	-0.962	0.336

Table 58. The overview of the generalized linear mixed-effects model run for the C1 post-hoc analysis.

Fixed effects	Estimate	SE	z-value	Pr(>|z|)
Intercept	-4.649	5.289	-0.879	0.379
MGL_R (similarity to regulars)	7.000	5.452	1.284	0.229
MGL_IR (similarity to irregulars)	13.845	12.298	1.926	0.075 .
MGL_R:MGL_IR	-15.530	12.896	-1.204	0.229

Table 59. The overview of the generalized linear mixed-effects model run for the post-hoc analysis of native speakers.

Fixed effects	Estimate	SE	z-value	Pr(>|z|)
Intercept	-3.005	7.746	-0.388	0.698
MGL_R (similarity to regulars)	28.371	12.963	3.549	0.016 **
MGL_IR (similarity to irregulars)	-7.402	5.599	-0.957	0.229
MGL_R:MGL_IR	37.348	20.605	1.813	0.069 .

Table 60. The overview of the linear mixed-effects model run for the initial interference analysis of reaction times. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

Fixed effects	Estimate	SE	t-value	Pr(>|t|)
Intercept	6.338e-01	5.099e-02	12.431	< .001 ***
LevelA2	-2.169e-02	6.486e-02	-0.334	0.738
LevelB1	-2.119e-03	6.194e-02	-0.034	0.972
LevelB2	-6.923e-02	6.121e-02	-1.131	0.260
LevelC1	-4.084e-02	6.046e-02	-0.675	0.500
LevelNS	-3.159e-02	6.215e-02	-0.852	0.378
SetB	7.057e-02	3.443e-02	1.049	0.422
Produced_form	-3.534e-02	1.877e-02	-1.992	0.069 .
LevelA2:SetB	-2.090e-02	3.735e-02	-0.560	0.575
LevelB1:SetB	-2.031e-02	3.534e-02	-1.917	0.072 .
LevelB2:SetB	-4.837e-02	3.485e-02	-1.388	0. 165
LevelC1:SetB	-4.808e-02	3.433e-02	-1.400	0.162
LevelNS:SetB	-4824e-02	3.479e-02	-1.412	0.154

Table 61. The overview of the linear mixed-effects model run for the analysis of the similarities to (ir)regulars.

Fixed effects	Estimate	SE	t-value	Pr(>|t|)
Intercept	0.596	0.372	1.602	0.119
MGL_R (similarity to regulars)	-1.794	0.899	-1.995	0.054 .
MGL_IR (similarity to irregulars)	-0.032	0.383	-0.084	0.933
MGL_R:MGL_IR	1.984	0.943	2.103	0.043 *

Appendix 3 – Experiment 3: The lexical decision experiment

Table 62. Novel and real-word stimuli used in Experiment 3 (novel words adopted from Experiment 2).

Novel words	
1	**bize**
2	**blafe**
3	**blig**
4	**bridge**
5	**chake**
6	**dape**
7	**dize**
8	**drice**
9	**drit**
10	**fledge**
11	**gare**
12	**gezz**
13	**gleed**
14	**glip**
15	**gude**
16	**nace**
17	**nold**
18	**nung**
19	**pank**
20	**plim**
21	**preak**
22	**rask**
23	**shilk**
24	**skride**
25	**spling**
26	**stin**
27	**stip**
28	**stire**
29	**teep**
30	**tesh**
31	**trisk**
32	**wiss**

	Real words		Frequency
most frequent	1	**know**	68,060
	2	**see**	56,997
	3	**think**	54,743
	4	**want**	31,141
	5	**mean**	29,017
	6	**get**	28,752
	7	**go**	26,403
	8	**say**	26,079
	9	**come**	23,865
	10	**need**	21,491
	11	**look**	20,151
medium frequent	12	**praise**	17
	13	**style**	17
	14	**wreak**	17
	15	**snore**	16
	16	**waive**	16
	17	**patch**	16
	18	**purge**	16
	19	**fine**	16
	20	**sport**	16
	21	**flake**	16
least frequent	22	**slurp**	5
	23	**lace**	5
	24	**ice**	5
	25	**bleat**	5
	26	**douse**	5
	27	**traipse**	5
	28	**shirk**	5
	29	**rail**	5
	30	**quell**	5
	31	**swipe**	5
	32	**chafe**	5

Table 63. The overview of the linear mixed-effects model run for the initial reaction-times analysis.

Fixed effects	**Estimate**	**SE**	**t-value**	**Pr(>|t|)**
Intercept	0.778	0.062	12.503	**< .001 *****
TypeRW (real words)	-0.122	0.042	-2.903	0.004 **
Score	-0.068	0.095	-0.716	0.478
TypeRW:Score	-0.026	0.055	-0.482	0.629

Table 64. The overview of the linear mixed-effects model run for the subsequent reaction-times analysis.

Fixed effects	**Estimate**	**SE**	**t-value**	**Pr(>|t|)**
Intercept	0.785	0.080	9.705	< .001 ***
Score	-0.100	0.124	-0.803	0.427
SetB	-0.021	0.060	-0.354	0.724
Score:SetB	0.084	0.085	0.986	0.325

Table 65. The overview of the generalized linear mixed-effects model run for the initial accuracy analysis.

Fixed effects	**Estimate**	**SE**	**z-value**	**Pr(>|z|)**
Intercept	1.278	0.331	3.865	< .001 ***
TypeRW (real words)	-0.736	0.459	-1.606	0.108
Score	-0.185	0.361	-0.511	0.609
TypeRW:Score	1.384	0.491	2.817	0.005 **

Table 66. The overview of the generalized linear mixed-effects model run for the subsequent accuracy analysis.

Fixed effects	**Estimate**	**SE**	**z-value**	**Pr(>|z|)**
Intercept	1.258	0.476	2.644	0.008 **
Score	0.245	0.695	0.353	0.724
SetB	0.146	0.495	0.296	0.767
Score:SetB	-0.908	0.661	-1.373	0.169

Appendix 4 – General discussion and conclusion

The lists of English regular and irregular verbs divided by structural analogy (practical application of the findings):

Table 67. Similarity groups of English regular verbs for the A1 level.

VERB	past tense	past participle	allomorph	vowel	syllables	class
cook	cooked	cooked	t	ʊ	1	D1t
look	looked	looked	t	ʊ	1	D1t
stop	stopped	stopped	t	ɒ	1	E1t
wash	washed	washed	t	ɒ	1	E1t
watch	watched	watched	t	ɒ	1	E1t
kick	kicked	kicked	t	ɪ	1	G1t
work	worked	worked	t	ɜ:	1	I1t
talk	talked	talked	t	ɔ:	1	J1t
walk	walked	walked	t	ɔ:	1	J1t
dance	danced	danced	t	ɑ:/æ	1	K/B1t
ask	asked	asked	t	ɑ:	1	K1t
like	liked	liked	t	aɪ	1	Q1t
smoke	smoked	smoked	t	əʊ	1	R1t
practise	practised	practised	t	æ	2	B2t
finish	finished	finished	t	ɪ	2	G2t
want	wanted	wanted	ɪd	ɒ	1	E1ɪd
need	needed	needed	ɪd	i:	1	H1ɪd
start	started	started	ɪd	ɑ:	1	L1ɪd
paint	painted	painted	ɪd	eɪ	1	O1ɪd
wait	waited	waited	ɪd	eɪ	1	O1ɪd
visit	visited	visited	ɪd	ɪ	2	G2ɪd
invite	invited	invited	ɪd	aɪ	2	Q2ɪd
help	helped	helped	d	e	1	A1d
love	loved	loved	d	ʌ	1	C1d
clean	cleaned	cleaned	d	i:	1	H1d
learn	learned	learned	d	ɜ:	1	I1d
use	used	used	d	u:	1	K1d
change	changed	changed	d	eɪ	1	O1d
pay	paid	paid	d	eɪ	1	O1d
play	played	played	d	eɪ	1	O1d
stay	stayed	stayed	d	eɪ	1	O1d
die	died	died	d	aɪ	1	Q1d
close	closed	closed	d	əʊ	1	R1d
phone	phoned	phoned	d	əʊ	1	R1d
carry	carried	carried	d	æ	2	B2d
travel	travelled	travelled	d	æ	2	B2d

colour	coloured	coloured	d	ʌ	2	C2d
study	studied	studied	d	ʌ	2	C2d
listen	listened	listened	d	ɪ	2	G2d
answer	answered	answered	d	ɑ:	2	L2d
enjoy	enjoyed	enjoyed	d	ɔɪ	2	P2d
open	opened	opened	d	əʊ	2	R2d
remember	remembered	remembered	d	e	3	A3d

Table 68. Similarity groups of English irregular verbs for the A1 level.

VERB	past tense	past participle	class	subcategory
learn	learnt	learnt	1	A
send	sent	sent	1	B
have	had	had	1	C
make	made	made	1	C
show	showed	shown	2	
feel	felt	felt	3	A
leave	left	left	3	A
sleep	slept	slept	3	A
buy	bought	bought	3	B
catch	caught	caught	3	B
teach	taught	taught	3	B
think	thought	thought	3	B
tell	told	told	3	D
hear	heard	heard	3	E
say	said	said	3	F
choose	chose	chosen	4	Aa
speak	spoke	spoken	4	Aa
wear	wore	worn	4	Ab
draw	drew	drawn	4	B
know	knew	known	4	Ba
take	took	taken	4	Bb
give	gave	given	4	Bc
eat	ate	eaten	4	Bf
see	saw	seen	4	Bg
do	did	done	4	C
ride	rode	ridden	4	Ca
write	wrote	written	4	Ca
fly	flew	flown	4	Cb
drive	drove	driven	4	Cc
meet	met	met	6	A
read	read	read	6	A
find	found	found	6	C
sit	sat	sat	6	E
get	got	got/gotten	6	F

understand	understood	understood	6	H
begin	began	begun	7	A
drink	drank	drunk	7	A
sing	sang	sung	7	A
swim	swam	swum	7	A
come	came	come	7	B
run	ran	run	7	B
go	went	gone	7	C
be	was/were	been	special	

Table 69. Similarity groups of English regular verbs for the A2 level.

VERB	past tense	past participle	allomorph	vowel	syllables	class
check	checked	checked	t	e	1	A1t
dress	dressed	dressed	t	e	1	A1t
guess	guessed	guessed	t	e	1	A1t
camp	camped	camped	t	æ	1	B1t
pack	packed	packed	t	æ	1	B1t
thank	thanked	thanked	t	æ	1	B1t
brush	brushed	brushed	t	ʌ	1	C1t
jump	jumped	jumped	t	ʌ	1	C1t
book	booked	booked	t	ʊ	1	D1t
push	pushed	pushed	t	ʊ	1	D1t
cross	crossed	crossed	t	ɒ	1	E1t
click	clicked	clicked	t	ɪ	1	G1t
kiss	kissed	kissed	t	ɪ	1	G1t
miss	missed	missed	t	ɪ	1	G1t
mix	mixed	mixed	t	ɪ	1	G1t
surf	surfed	surfed	t	ɜ:	1	I1t
laugh	laughed	laughed	t	ɑ:	1	L1t
park	parked	parked	t	ɑ:	1	L1t
pass	passed	passed	t	ɑ:	1	L1t
bake	baked	baked	t	eɪ	1	O1t
hope	hoped	hoped	t	əʊ	1	R1t
discuss	discussed	discussed	t	ʌ	2	C2t
end	ended	ended	ɪd	e	1	A1ɪd
rent	rented	rented	ɪd	e	1	A1ɪd
text	texted	texted	ɪd	e	1	A1ɪd
add	added	added	ɪd	æ	1	B1ɪd
chat	chatted	chatted	ɪd	æ	1	B1ɪd
print	printed	printed	ɪd	ɪ	1	G1ɪd
hate	hated	hated	ɪd	eɪ	1	O1ɪd
point	pointed	pointed	ɪd	ɔɪ	1	P1ɪd
post	posted	posted	ɪd	əʊ	1	R1ɪd
roast	roasted	roasted	ɪd	əʊ	1	R1ɪd

shout	shouted	shouted	ɪd	aʊ	1	S1ɪd
collect	collected	collected	ɪd	e	2	A2ɪd
contact	contacted	contacted	ɪd	æ	2	B2ɪd
complete	completed	completed	ɪd	i:	2	H2ɪd
repeat	repeated	repeated	ɪd	i:	2	H2ɪd
record	recorded	recorded	ɪd	ɔ:	2	J2ɪd
include	include	included	ɪd	u:	2	K2ɪd
decide	decided	decided	ɪd	aɪ	2	Q2ɪd
download	downloaded	downloaded	ɪd	aʊ	2	S2ɪd
spell	spelled	spelled	d	e	1	A1D
pull	pulled	pulled	d	ʊ	1	D1d
fill	filled	filled	d	ɪ	1	G1d
kill	killed	killed	d	ɪ	1	G1d
dream	dreamed	dreamed	d	i:	1	H1d
earn	earned	earned	d	ɜ:	1	I1d
serve	served	served	d	ɜ:	1	I1d
turn	turned	turned	d	ɜ:	1	I1d
call	called	called	d	ɔ:	1	J1d
move	moved	moved	d	u:	1	K1d
share	shared	shared	d	eə	1	N1d
fail	failed	failed	d	eɪ	1	O1d
pray	prayed	prayed	d	eɪ	1	O1d
save	saved	saved	d	eɪ	1	O1d
boil	boiled	boiled	d	ɔɪ	1	P1d
join	joined	joined	d	ɔɪ	1	P1d
climb	climbed	climbed	d	aɪ	1	Q1d
cry	cried	cried	d	aɪ	1	Q1d
dry	dried	dried	d	aɪ	1	Q1d
try	tried	tried	d	aɪ	1	Q1d
snow	snowed	snowed	d	əʊ	1	R1D
enter	entered	entered	d	e	2	A2d
happen	happened	happened	d	æ	2	B2d
matter	mattered	mattered	d	æ	2	B2d
cover	covered	covered	d	ʌ	2	C2d
hurry	hurried	hurried	d	ʌ	2	C2d
worry	worried	worried	d	ʌ	2	C2d
offer	offered	offered	d	ɒ	2	E2d
borrow	borrowed	borrowed	d	ɒ	2	E2d
bother	bothered	bothered	d	ɒ	2	E2d
copy	copied	copied	d	ɒ	2	E2d
follow	followed	followed	d	ɒ	2	E2d
agree	agreed	agreed	d	i:	2	H2d
believe	believed	believed	d	i:	2	H2d
email	emailed	emailed	d	i:	2	H2d
receive	received	received	d	i:	2	H2d
prefer	preferred	preferred	d	ɜ:	2	I2d
return	returned	returned	d	ɜ:	2	I2d

order	ordered	ordered	d	ɔ:	2	J2d
improve	improved	improved	d	u:	2	K2d
prepare	prepared	prepared	d	eə	2	N2d
repair	repaired	repaired	d	eə	2	N2d
delay	delayed	delayed	d	eɪ	2	O2d
explain	explained	explained	d	eɪ	2	O2d
arrive	arrived	arrived	d	aɪ	2	Q2d
describe	described	described	d	aɪ	2	Q2d
tidy	tidied	tidied	d	aɪ	2	Q2d
telephone	telephoned	telephoned	d	e	3	A3d

Table 70. Similarity groups of English irregular verbs for the A2 level.

VERB	past tense	past participle	class	subcategory
build	built	built	1	B
lend	lent	lent	1	B
spend	spent	spent	1	B
dream	dreamt	dreamt	3	A
keep	kept	kept	3	A
mean	meant	meant	3	A
bring	brought	brought	3	B
lose	lost	lost	3	C
sell	sold	sold	3	D
break	broke	broken	4	Aa
steal	stole	stolen	4	Aa
forget	forgot	forgotten	4	Ad
lie	lay	lain	4	Ae
grow	grew	grown	4	Ba
throw	threw	thrown	4	Ba
fall	fell	fallen	4	Be
cost	cost	cost	5	
cut	cut	cut	5	
hit	hit	hit	5	
hurt	hurt	hurt	5	
put	put	put	5	
shut	shut	shut	5	
hold	held	held	6	A
win	won	won	6	B
stand	stood	stood	6	H
ring	rang	rung	7	A

Table 71. Similarity groups of English regular verbs for the B1 level.

VERB	past tense	past participle	allomorph	vowel	syllables	class
fetch	fetched	fetched	t	e	1	A1t
press	pressed	pressed	t	e	1	A1t
clap	clapped	clapped	t	æ	1	B1t
crash	crashed	crashed	t	æ	1	B1t
fax	faxed	faxed	t	æ	1	B1t
match	matched	matched	t	æ	1	B1t
wrap	wrapped	wrapped	t	æ	1	B1t
touch	touched	touched	t	ʌ	1	C1t
cough	coughed	coughed	t	ɒ	1	E1t
drop	dropped	dropped	t	ɒ	1	E1t
knock	knocked	knocked	t	ɒ	1	E1t
lock	locked	locked	t	ɒ	1	E1t
shop	shopped	shopped	t	ɒ	1	E1t
fish	fished	fished	t	ɪ	1	G1t
fix	fixed	fixed	t	ɪ	1	G1t
pick	picked	picked	t	ɪ	1	G1t
slip	slipped	slipped	t	ɪ	1	G1t
stick	sticked	sticked	t	ɪ	1	G1t
tick	ticked	ticked	t	ɪ	1	G1t
reach	reached	reached	t	i:	1	H1t
search	searched	searched	t	ɜ:	1	I1t
face	faced	faced	t	eɪ	1	O1t
race	raced	raced	t	eɪ	1	O1t
wake	waked	waked	t	eɪ	1	O1t
type	typed	typed	t	aɪ	1	Q1t
joke	joked	joked	t	əʊ	1	R1t
undress	undressed	undressed	t	e	2	A3t
attach	attached	attached	t	æ	2	B2t
attack	attacked	attacked	t	æ	2	B2t
relax	relaxed	relaxed	t	æ	2	B2t
unpack	unpacked	unpacked	t	æ	2	B2t
publish	published	published	t	ʌ	2	C1t
punish	punished	punished	t	ʌ	2	C1t
promise	promised	promised	t	ɒ	2	E2t
convince	convinced	convinced	t	ɪ	2	G2t
hitchhike	hitchhiked	hitchhiked	t	ɪ	2	G2t
decrease	decreased	decreased	t	i:	2	H2t
increase	increased	increased	t	i:	2	H2t
produce	produced	produced	t	u:	2	K2t
reduce	reduced	reduced	t	u:	2	K2t
escape	escaped	escaped	t	eɪ	2	O2t
replace	replaced	replaced	t	eɪ	2	O2t
dislike	disliked	disliked	t	aɪ	2	Q2t

approach	approached	approached	t	əʊ	2	R2t
notice	noticed	noticed	t	əʊ	2	R2t
announce	announced	announced	t	aʊ	2	S2t
pronounce	pronounced	pronounced	t	aʊ	2	S2t
develop	developed	developed	t	e	3	A3t
introduce	introduced	introduced	t	u:	3	K3t
experience	experienced	experienced	t	ɪə	4	M4t
mend	mended	mended	ɪd	e	1	A1ɪd
rest	rested	rested	ɪd	e	1	A1ɪd
test	tested	tested	ɪd	e	1	A1ɪd
act	acted	acted	ɪd	æ	1	B1ɪd
hand	handed	handed	ɪd	æ	1	B1ɪd
land	landed	landed	ɪd	æ	1	B1ɪd
flood	flooded	flooded	ɪd	ʌ	1	C1ɪd
hunt	hunted	hunted	ɪd	ʌ	1	C1ɪd
trust	trusted	trusted	ɪd	ʌ	1	C1ɪd
fit	fitted	fitted	ɪd	ɪ	1	G1ɪd
knit	knitted	knitted	ɪd	ɪ	1	G1ɪd
lift	lifted	lifted	ɪd	ɪ	1	G1ɪd
quit	quitted	quitted	ɪd	ɪ	1	G1ɪd
greet	greeted	greeted	ɪd	i:	1	H1ɪd
board	boarded	boarded	ɪd	ɔ:	1	J1ɪd
last	lasted	lasted	ɪd	ɑ:	1	L1ɪd
plant	planted	planted	ɪd	ɑ:/æ	1	L1ɪd/B1ɪd
date	dated	dated	ɪd	eɪ	1	O1ɪd
skate	skated	skated	ɪd	eɪ	1	O1ɪd
taste	tasted	tasted	ɪd	eɪ	1	O1ɪd
waste	wasted	wasted	ɪd	eɪ	1	O1ɪd
guide	guided	guided	ɪd	aɪ	1	Q1ɪd
light	lighted	lighted	ɪd	aɪ	1	Q1ɪd
mind	minded	minded	ɪd	aɪ	1	Q1ɪd
float	floated	floated	ɪd	əʊ	1	R1ɪd
fold	folded	folded	ɪd	əʊ	1	R1ɪd
note	noted	noted	ɪd	əʊ	1	R1ɪd
vote	voted	voted	ɪd	əʊ	1	R1ɪd
count	counted	counted	ɪd	aʊ	1	S1ɪd
accept	accepted	accepted	ɪd	e	2	A2ɪd
arrest	arrested	arrested	ɪd	e	2	A2ɪd
attempt	attempted	attempted	ɪd	e	2	A2ɪd
attend	attended	attended	ɪd	e	2	A2ɪd
connect	connected	connected	ɪd	e	2	A2ɪd
defend	defended	defended	ɪd	e	2	A2ɪd
direct	directed	directed	ɪd	e	2	A2ɪd
exit	exited	exited	ɪd	e	2	A2ɪd
expect	expected	expected	ɪd	e	2	A2ɪd
intend	intended	intended	ɪd	e	2	A2ɪd
invent	invented	invented	ɪd	e	2	A2ɪd

prevent	prevented	prevented	ɪd	e	2	A2ɪd
protect	protected	protected	ɪd	e	2	A2ɪd
regret	regretted	regretted	ɪd	e	2	A2ɪd
request	requested	requested	ɪd	e	2	A2ɪd
respect	respected	respected	ɪd	e	2	A2ɪd
select	selected	selected	ɪd	e	2	A2ɪd
suggest	suggested	suggested	ɪd	e	2	A2ɪd
attract	attracted	attracted	ɪd	æ	2	B2ɪd
admit	admitted	admitted	ɪd	ɪ	2	G2ɪd
exist	existed	existed	ɪd	ɪ	2	G2ɪd
insist	insisted	insisted	ɪd	ɪ	2	G2ɪd
permit	permitted	permitted	ɪd	ɪ	2	G2ɪd
predict	predicted	predicted	ɪd	ɪ	2	G2ɪd
compete	competed	competed	ɪd	i:	2	H2ɪd
defeat	defeated	defeated	ɪd	i:	2	H2ɪd
delete	deleted	deleted	ɪd	i:	2	H2ɪd
succeed	succeeded	succeeded	ɪd	i:	2	H2ɪd
report	reported	reported	ɪd	ɔ:	2	J2ɪd
support	supported	supported	ɪd	ɔ:	2	J2ɪd
demand	demanded	demanded	ɪd	ɑ:	2	L2ɪd
depart	departed	departed	ɪd	ɑ:	2	L2ɪd
create	created	created	ɪd	eɪ	2	O2ɪd
persuade	persuaded	persuaded	ɪd	eɪ	2	O2ɪd
translate	translated	translated	ɪd	eɪ	2	O2ɪd
update	updated	updated	ɪd	eɪ	2	O2ɪd
avoid	avoided	avoided	ɪd	ɔɪ	2	P2ɪd
divide	divided	divided	ɪd	aɪ	2	Q1ɪd
provide	provided	provided	ɪd	aɪ	2	Q1ɪd
remind	reminded	reminded	ɪd	aɪ	2	Q1ɪd
explode	exploded	exploded	ɪd	əʊ	2	R2ɪd
promote	promoted	promoted	ɪd	əʊ	2	R2ɪd
upload	uploaded	uploaded	ɪd	əʊ	2	R2ɪd
surround	surrounded	surrounded	ɪd	aʊ	2	S2ɪd
celebrate	celebrated	celebrated	ɪd	e	3	A3ɪd
decorate	decorated	decorated	ɪd	e	3	A3ɪd
recommend	recommended	recommended	ɪd	e	3	A3ɪd
interrupt	interrupted	interrupted	ɪd	ʌ	3	C3ɪd
concentrate	concentrated	concentrated	ɪd	ɒ	3	E3ɪd
operate	operated	operated	ɪd	ɒ	3	E3ɪd
interest	interested	interested	ɪd	ɪ	3	G3ɪd
disappoint	disappointed	disappointed	ɪd	ɔɪ	3	P3ɪd
communicate	communicated	communicated	ɪd	u:	4	K4ɪd
smell	smelled	smelled	d	e	1	A1d
drag	dragged	dragged	d	æ	1	B1d
grab	grabbed	grabbed	d	æ	1	B1d
hang	hanged	hanged	d	æ	1	B1d
plan	planned	planned	d	æ	1	B1d

hug	hugged	hugged	d	ʌ	1	C1d
judge	judged	judged	d	ʌ	1	C1d
blog	blogged	blogged	d	ɒ	1	E1d
bomb	bombed	bombed	d	ɒ	1	E1d
jog	jogged	jogged	d	ɒ	1	E1d
rob	robbed	robbed	d	ɒ	1	E1d
solve	solved	solved	d	ɒ	1	E1d
film	filmed	filmed	d	ɪ	1	G1d
grill	grilled	grilled	d	ɪ	1	G1d
live	lived	lived	d	ɪ	1	G1d
pin	pinned	pinned	d	ɪ	1	G1d
spill	spilled	spilled	d	ɪ	1	G1d
ski	skied	skied	d	i:	1	H1d
breathe	breathed	breathed	d	i:	1	H1d
please	pleased	pleased	d	i:	1	H1d
scream	screamed	screamed	d	i:	1	H1d
burn	burned	burned	d	ɜ:	1	I1d
stir	stirred	stirred	d	ɜ:	1	I1d
form	formed	formed	d	ɔ:	1	J1d
pause	paused	paused	d	ɔ:	1	J1d
pour	poured	poured	d	ɔ:	1	J1d
score	scored	scored	d	ɔ:	1	J1d
warn	warned	warned	d	ɔ:	1	J1d
prove	proved	proved	d	u:	1	K1d
charge	charged	charged	d	ɑ:	1	L1d
star	stared	stared	d	ɑ:	1	L1d
clear	cleared	cleared	d	ɪə	1	M1d
care	cared	cared	d	eə	1	N1d
blame	blamed	blamed	d	eɪ	1	O1d
gain	gained	gained	d	eɪ	1	O1d
name	named	named	d	eɪ	1	O1d
raise	raised	raised	d	eɪ	1	O1d
sail	sailed	sailed	d	eɪ	1	O1d
shave	shaved	shaved	d	eɪ	1	O1d
train	trained	trained	d	eɪ	1	O1d
wave	waved	waved	d	eɪ	1	O1d
weigh	weighed	weighed	d	eɪ	1	O1d
spoil	spoiled	spoiled	d	ɔɪ	1	P1d
dive	dived	dived	d	aɪ	1	Q1d
lie	lied	lied	d	aɪ	1	Q1d
shine	shined	shined	d	aɪ	1	Q1d
sign	signed	signed	d	aɪ	1	Q1d
smile	smiled	smiled	d	aɪ	1	Q1d
tie	tied	tied	d	aɪ	1	Q1d
fry	fried	fried	d	aɪ	1	Q1d
comb	combed	combed	d	əʊ	1	R1d
flow	flowed	flowed	d	əʊ	1	R1d

owe	owed	owed	d	əʊ	1	R1d
own	owned	owned	d	əʊ	1	R1d
sew	sewed	sewed	d	əʊ	1	R1d
tour	toured	toured	d	ʊə	1	T1d
hire	hired	hired	d	aɪə	1	W1d
bury	buried	buried	d	e	2	A2d
mention	mentioned	mentioned	d	e	2	A2d
rescue	rescued	rescued	d	e	2	A2d
welcome	welcomed	welcomed	d	e	2	A2d
fancy	fancied	fancied	d	æ	2	B2d
cancel	cancelled	cancelled	d	æ	2	B2d
damage	damaged	damaged	d	æ	2	B2d
handle	handled	handled	d	æ	2	B2d
manage	managed	managed	d	æ	2	B2d
marry	married	married	d	æ	2	B2d
sunbathe	sunbathed	sunbathed	d	ʌ	2	C2d
wonder	wondered	wondered	d	ʌ	2	C2d
involve	involved	involved	d	ɒ	2	E2d
injure	injured	injured	d	ɪ	2	G2d
achieve	achieved	achieved	d	i:	2	H2d
circle	circled	circled	d	ɜ:	2	I2d
confirm	confirmed	confirmed	d	ɜ:	2	I2d
deserve	deserved	deserved	d	ɜ:	2	I2d
murder	murdered	murdered	d	ɜ:	2	I2d
reserve	reserved	reserved	d	ɜ:	2	I2d
explore	explored	explored	d	ɔ:	2	J2d
inform	informed	informed	d	ɔ:	2	J2d
install	installed	installed	d	ɔ:	2	J2d
perform	performed	performed	d	ɔ:	2	J2d
approve	approved	approved	d	u:	2	K2d
excuse	excused	excused	d	u:	2	K2d
refuse	refused	refused	d	u:	2	K2d
remove	removed	removed	d	u:	2	K2d
argue	argued	argued	d	ɑ:	2	L2d
fasten	fastened	fastened	d	ɑ:	2	L2d
transfer	transferred	transferred	d	ɑ:/æ	2	L2d/B2d
appear	appeared	appeared	d	ɪə	2	M2d
compare	compared	compared	d	eə	2	N2d
declare	declared	declared	d	eə	2	N2d
arrange	arranged	arranged	d	eɪ	2	O2d
behave	behaved	behaved	d	eɪ	2	O2d
complain	complained	complained	d	eɪ	2	O2d
contain	contained	contained	d	eɪ	2	O2d
display	displayed	displayed	d	eɪ	2	O2d
exchange	exchanged	exchanged	d	eɪ	2	O2d
annoy	annoyed	annoyed	d	ɔɪ	2	P2d
destroy	destroyed	destroyed	d	ɔɪ	2	P2d

employ	employed	employed	d	ɔɪ	2	P2d
design	designed	designed	d	aɪ	2	Q2d
advise	advised	advised	d	aɪ	2	Q2d
apply	applied	applied	d	aɪ	2	Q2d
cycle	cycled	cycled	d	aɪ	2	Q2d
dial	dialled	dialled	d	aɪ	2	Q2d
iron	ironed	ironed	d	aɪ	2	Q2d
realize	realized	realized	d	aɪ	2	Q2d
reply	replied	replied	d	aɪ	2	Q2d
revise	revised	revised	d	aɪ	2	Q2d
surprise	surprised	surprised	d	aɪ	2	Q2d
control	controlled	controlled	d	əʊ	2	R2d
postpone	postponed	postponed	d	əʊ	2	R2d
allow	allowed	allowed	d	aʊ	2	S2d
admire	admired	admired	d	aɪə	2	W2d
require	required	required	d	aɪə	2	W2d
retire	retired	retired	d	aɪə	2	W2d
exercise	exercised	exercised	d	e	3	A3d
recognize	recognized	recognized	d	e	3	A3d
register	registered	registered	d	e	3	A3d
advertise	advertised	advertised	d	æ	3	B3d
imagine	imagined	imagined	d	æ	3	B3d
discover	discovered	discovered	d	ʌ	3	C3d
recover	recovered	recovered	d	ʌ	3	C3d
encourage	encouraged	encouraged	d	ʌ	3	C3d
underline	underlined	underlined	d	ʌ	3	C3d
deliver	delivered	delivered	d	ɪ	3	G3d
consider	considered	considered	d	ɪ	3	G3d
continue	continued	continued	d	ɪ	3	G3d
interview	interviewed	interviewed	d	ɪ	3	G3d
disagree	disagreed	disagreed	d	i:	3	H3d
organize	organized	organized	d	ɔ:	3	J3d
barbecue	barbecued	barbecued	d	ɑ:	3	L2d
disappear	disappeared	disappeared	d	ɪə	3	M3d
entertain	entertained	entertained	d	eɪ	3	O3d
recycle	recycled	recycled	d	aɪ	3	Q3d
accompany	accompanied	accompanied	d	ʌ	4	C4d
apologise	apologised	apologised	d	ɒ	4	E4d

Table 72. Similarity groups of English irregular verbs for the B1 level.

VERB	past tense	past participle	class	subcategory
burn	burnt	burnt	1	A
rebuild	rebuilt	rebuilt	1	B
sew	sewed	sewn	2	
freeze	froze	frozen	4	Aa

wake	woke	woken	4	Aa
tear	tore	torn	4	Ab
bite	bit	bit/bitten	4	Ac
hide	hid	hid/hidden	4	Ac
blow	blew	blown	4	Ba
shake	shook	shaken	4	Bb
forgive	forgave	forgiven	4	Bc
rise	rose	risen	4	Ca
beat	beat	beat/beaten	4	D
dive	dove	dived	4	E
fit	fit	fit	5	
knit	knit	knit	5	
let	let	let	5	
quit	quit	quit	5	
set	set	set	5	
bleed	bled	bled	6	A
feed	fed	fed	6	A
lead	led	led	6	A
dig	dug	dug	6	B
hang	hung	hung	6	B
strike	struck	struck	6	B
light	lit	lit	6	D
babysit	babysat	babysat	6	E
shine	shone	shone	6	F
shoot	shot	shot	6	F
fight	fought	fought	6	G
sink	sank/sunk	sunk	7	A

Table 73. Similarity groups of English regular verbs for the B2 level.

VERB	past tense	past participle	allomorph	vowel	syllables	class
stress	stressed	stressed	t	e	1	A1t
stretch	stretched	stretched	t	e	1	A1t
crack	cracked	cracked	t	æ	1	B1t
dash	dashed	dashed	t	æ	1	B1t
flash	flashed	flashed	t	æ	1	B1t
lack	lacked	lacked	t	æ	1	B1t
sack	sacked	sacked	t	æ	1	B1t
scratch	scratched	scratched	t	æ	1	B1t
slap	slapped	slapped	t	æ	1	B1t
smash	smashed	smashed	t	æ	1	B1t
splash	splashed	splashed	t	æ	1	B1t
stamp	stamped	stamped	t	æ	1	B1t
tap	tapped	tapped	t	æ	1	B1t
blush	blushed	blushed	t	ʌ	1	C1t

bump	bumped	bumped	t	ʌ	1	C1t
punch	punched	punched	t	ʌ	1	C1t
rush	rushed	rushed	t	ʌ	1	C1t
block	blocked	blocked	t	ɒ	1	E1t
chop	chopped	chopped	t	ɒ	1	E1t
shock	shocked	shocked	t	ɒ	1	E1t
squash	squashed	squashed	t	ɒ	1	E1t
blink	blinked	blinked	t	ɪ	1	G1t
dip	dipped	dipped	t	ɪ	1	G1t
grip	gripped	gripped	t	ɪ	1	G1t
itch	itched	itched	t	ɪ	1	G1t
lick	licked	licked	t	ɪ	1	G1t
link	linked	linked	t	ɪ	1	G1t
rip	ripped	ripped	t	ɪ	1	G1t
risk	risked	risked	t	ɪ	1	G1t
skip	skipped	skipped	t	ɪ	1	G1t
switch	switched	switched	t	ɪ	1	G1t
trick	tricked	tricked	t	ɪ	1	G1t
trip	tripped	tripped	t	ɪ	1	G1t
cease	ceased	ceased	t	iː	1	H1t
leak	leaked	leaked	t	iː	1	H1t
force	forced	forced	t	ɔː	1	J1t
launch	launched	launched	t	ɔː	1	J1t
bark	barked	barked	t	ɑː	1	L1t
mark	marked	marked	t	ɑː	1	L1t
ache	ached	ached	t	eɪ	1	O1t
brake	braked	braked	t	eɪ	1	O1t
chase	chased	chased	t	eɪ	1	O1t
rape	raped	raped	t	eɪ	1	O1t
slice	sliced	sliced	t	aɪ	1	Q1t
wipe	wiped	wiped	t	aɪ	1	Q1t
coach	coached	coached	t	əʊ	1	R1t
cope	coped	coped	t	əʊ	1	R1t
soak	soaked	soaked	t	əʊ	1	R1t
stroke	stroked	stroked	t	əʊ	1	R1t
bounce	bounced	bounced	t	aʊ	1	S1t
assess	assessed	assessed	t	e	2	A2t
confess	confessed	confessed	t	e	2	A2t
express	expressed	expressed	t	e	2	A2t
sentence	sentenced	sentenced	t	e	2	A2t
access	accessed	accessed	t	æ	2	B2t
balance	balanced	balanced	t	æ	2	B2t
collapse	collapsed	collapsed	t	æ	2	B2t
finance	financed	financed	t	æ	2	B2t
panic	panicked	panicked	t	æ	2	B2t
vanish	vanished	vanished	t	æ	2	B2t
bookmark	bookmarked	bookmarked	t	ʊ	2	D2t

gossip	gossiped	gossiped	t	ɒ	2	E1t
unlock	unlocked	unlocked	t	ɒ	2	E2t
equip	equipped	equipped	t	ɪ	2	G2t
impress	impressed	impressed	t	ɪ	2	G2t
witness	witnessed	witnessed	t	ɪ	2	G2t
release	released	released	t	iː	2	H2t
purchase	purchased	purchased	t	ɜː	2	I2t
research	researched	researched	t	ɜː	2	I2t
reverse	reversed	reversed	t	ɜː	2	I2t
divorce	divorced	divorced	t	ɔː	2	J2t
remark	remarked	remarked	t	ɑː	2	L2t
process	processed	processed	t	əʊ	2	R2t
progress	progressed	progressed	t	əʊ	2	R2t
establish	established	established	t	æ	3	B3t
overlook	overlooked	overlooked	t	ʊ	3	D3t
abolish	abolished	abolished	t	ɒ	3	E3t
demolish	demolished	demolished	t	ɒ	3	E3t
distinguish	distinguished	distinguished	t	ɪ	3	G3t
influence	influenced	influenced	t	ɪ	3	G3t
photograph	photographed	photographed	t	əʊ	3	R3t
head	headed	headed	ɪd	e	1	A1ɪd
melt	melted	melted	ɪd	e	1	A1ɪd
sweat	sweated	sweated	ɪd	e	1	A1ɪd
nod	nodded	nodded	ɪd	ɒ	1	E1ɪd
spot	spotted	spotted	ɪd	ɒ	1	E1ɪd
list	listed	listed	ɪd	ɪ	1	G1ɪd
cheat	cheated	cheated	ɪd	iː	1	H1ɪd
treat	treated	treated	ɪd	iː	1	H1ɪd
haunt	haunted	haunted	ɪd	ɔː	1	J1ɪd
sort	sorted	sorted	ɪd	ɔː	1	J1ɪd
boost	boosted	boosted	ɪd	uː	1	K1ɪd
suit	suited	suited	ɪd	uː	1	K1ɪd
wound	wounded	wounded	ɪd	uː	1	K1ɪd
grant	granted	granted	ɪd	ɑː	1	L1ɪd
guard	guarded	guarded	ɪd	ɑː	1	L1ɪd
fade	faded	faded	ɪd	eɪ	1	O1ɪd
faint	fainted	fainted	ɪd	eɪ	1	O1ɪd
state	stated	stated	ɪd	eɪ	1	O1ɪd
trade	traded	traded	ɪd	eɪ	1	O1ɪd
boast	boasted	boasted	ɪd	əʊ	1	R1ɪd
load	loaded	loaded	ɪd	əʊ	1	R1ɪd
doubt	doubted	doubted	ɪd	aʊ	1	S1ɪd
found	founded	founded	ɪd	aʊ	1	S1ɪd
pound	pounded	pounded	ɪd	aʊ	1	S1ɪd
affect	affected	affected	ɪd	e	2	A2ɪd
correct	corrected	corrected	ɪd	e	2	A2ɪd
debit	debited	debited	ɪd	e	2	A2ɪd

descend	descended	descended	ɪd	e	2	A2ɪd
edit	edited	edited	ɪd	e	2	A2ɪd
elect	elected	elected	ɪd	e	2	A2ɪd
extend	extended	extended	ɪd	e	2	A2ɪd
invest	invested	invested	ɪd	e	2	A2ɪd
offend	offended	offended	ɪd	e	2	A2ɪd
present	presented	presented	ɪd	e	2	A2ɪd
pretend	pretended	pretended	ɪd	e	2	A2ɪd
protest	protested	protested	ɪd	e	2	A2ɪd
reflect	reflected	reflected	ɪd	e	2	A2ɪd
reject	rejected	rejected	ɪd	e	2	A2ɪd
suspect	suspected	suspected	ɪd	e	2	A2ɪd
suspend	suspended	suspended	ɪd	e	2	A2ɪd
adapt	adapted	adapted	ɪd	æ	2	B2ɪd
distract	distracted	distracted	ɪd	æ	2	B2ɪd
expand	expanded	expanded	ɪd	æ	2	B2ɪd
extract	extracted	extracted	ɪd	æ	2	B2ɪd
react	reacted	reacted	ɪd	æ	2	B2ɪd
adjust	adjusted	adjusted	ɪd	ʌ	2	C2ɪd
construct	constructed	constructed	ɪd	ʌ	2	C2ɪd
disrupt	disrupted	disrupted	ɪd	ʌ	2	C2ɪd
insult	insulted	insulted	ɪd	ʌ	2	C2ɪd
adopt	adopted	adopted	ɪd	ɒ	2	E2ɪd
comment	commented	commented	ɪd	ɒ	2	E2ɪd
conduct	conducted	conducted	ɪd	ɒ	2	E2ɪd
respond	responded	responded	ɪd	ɒ	2	E2ɪd
object	objected	objected	ɪd	ɪ/e	2	G/A2ɪd
assist	assisted	assisted	ɪd	ɪ	2	G2ɪd
commit	committed	committed	ɪd	ɪ	2	G2ɪd
limit	limited	limited	ɪd	ɪ	2	G2ɪd
resist	resisted	resisted	ɪd	ɪ	2	G2ɪd
submit	submitted	submitted	ɪd	ɪ	2	G2ɪd
convert	converted	converted	ɪd	ɜː	2	I2ɪd
award	awarded	awarded	ɪd	ɔː	2	J2ɪd
broadcast	broadcasted	broadcasted	ɪd	ɔː	2	J2ɪd
export	exported	exported	ɪd	ɔː	2	J2ɪd
import	imported	imported	ɪd	ɔː	2	J2ɪd
reward	rewarded	rewarded	ɪd	ɔː	2	J2ɪd
transport	transported	transported	ɪd	ɔː	2	J2ɪd
pollute	polluted	polluted	ɪd	uː	2	K2ɪd
regard	regarded	regarded	ɪd	ɑː	2	L2ɪd
donate	donated	donated	ɪd	eɪ	2	O2ɪd
invade	invaded	invaded	ɪd	eɪ	2	O2ɪd
locate	located	located	ɪd	eɪ	2	O2ɪd
upgrade	upgraded	upgraded	ɪd	eɪ	2	O2ɪd
exploit	exploited	exploited	ɪd	ɔɪ	2	P2ɪd
highlight	highlighted	highlighted	ɪd	aɪ	2	Q2ɪd

benefit	benefited	benefited	ɪd	e	3	A3ɪd
demonstrate	demonstrated	demonstrated	ɪd	e	3	A3ɪd
educate	educated	educated	ɪd	e	3	A3ɪd
estimate	estimated	estimated	ɪd	e	3	A3ɪd
generate	generated	generated	ɪd	e	3	A3ɪd
hesitate	hesitated	hesitated	ɪd	e	3	A3ɪd
represent	represented	represented	ɪd	e	3	A3ɪd
separate	separated	separated	ɪd	e	3	A3ɪd
calculate	calculated	calculated	ɪd	æ	3	B2ɪd
graduate	graduated	graduated	ɪd	æ	3	B3ɪd
interact	interacted	interacted	ɪd	æ	3	B3ɪd
substitute	substituted	substituted	ɪd	ʌ	3	C3ɪd
correspond	corresponded	corresponded	ɪd	ɒ	3	E3ɪd
dominate	dominated	dominated	ɪd	ɒ	3	E3ɪd
tolerate	tolerated	tolerated	ɪd	ɒ	3	E3ɪd
contribute	contributed	contributed	ɪd	ɪ	3	G3ɪd
distribute	distributed	distributed	ɪd	ɪ	3	G3ɪd
illustrate	illustrated	illustrated	ɪd	ɪ	3	G3ɪd
implement	implemented	implemented	ɪd	ɪ	3	G3ɪd
indicate	indicated	indicated	ɪd	ɪ	3	G3ɪd
prohibit	prohibited	prohibited	ɪd	ɪ	3	G3ɪd
stimulate	stimulated	stimulated	ɪd	ɪ	3	G3ɪd
interpret	interpreted	interpreted	ɪd	ɜː	3	I3ɪd
experiment	experimented	experimented	ɪd	e	4	A4ɪd
investigate	investigated	investigated	ɪd	e	4	A4ɪd
congratulate	congratulated	congratulated	ɪd	æ	4	B4ɪd
cooperate	cooperated	cooperated	ɪd	ɒ	4	E4ɪd
participate	participated	participated	ɪd	ɪ	4	G4ɪd
appreciate	appreciated	appreciated	ɪd	i:	4	H4ɪd
underestimate	underestimated	underestimated	ɪd	e	5	A5ɪd
beg	begged	begged	d	e	1	A1d
yell	yelled	yelled	d	e	1	A1d
ban	banned	banned	d	æ	1	B1d
bang	banged	banged	d	æ	1	B1d
slam	slammed	slammed	d	æ	1	B1d
stab	stabbed	stabbed	d	æ	1	B1d
rub	rubbed	rubbed	d	ʌ	1	C1d
sob	sobbed	sobbed	d	ɒ	1	E1d
swing	swinged	swinged	d	ɪ	1	G1d
free	freed	freed	d	i:	1	H1d
heal	healed	healed	d	i:	1	H1d
kneel	kneeled	kneeled	d	i:	1	H1d
peel	peeled	peeled	d	i:	1	H1d
seize	seized	seized	d	i:	1	H1d
sneeze	sneezed	sneezed	d	i:	1	H1d
squeeze	squeezed	squeezed	d	i:	1	H1d
tease	teased	teased	d	i:	1	H1d

cause	caused	caused	d	ɔː	1	J1d
crawl	crawled	crawled	d	ɔː	1	J1d
snore	snored	snored	d	ɔː	1	J1d
store	stored	stored	d	ɔː	1	J1d
warm	warmed	warmed	d	ɔː	1	J1d
yawn	yawned	yawned	d	ɔː	1	J1d
chew	chewed	chewed	d	uː	1	K1d
cool	cooled	cooled	d	uː	1	K1d
cruise	cruised	cruised	d	uː	1	K1d
fool	fooled	fooled	d	uː	1	K1d
queue	queued	queued	d	uː	1	K1d
rule	ruled	ruled	d	uː	1	K1d
harm	harmed	harmed	d	ɑː	1	L1d
cheer	cheered	cheered	d	ɪə	1	M1d
fear	feared	feared	d	ɪə	1	M1d
steer	steered	steered	d	ɪə	1	M1d
stare	stared	stared	d	eə	1	N1d
claim	claimed	claimed	d	eɪ	1	O1d
praise	praised	praised	d	eɪ	1	O1d
spray	sprayed	sprayed	d	eɪ	1	O1d
strain	strained	strained	d	eɪ	1	O1d
fine	fined	fined	d	aɪ	1	Q1d
sigh	signed	signed	d	aɪ	1	Q1d
time	timed	timed	d	aɪ	1	Q1d
roll	rolled	rolled	d	əʊ	1	R1d
row	rowed	rowed	d	əʊ	1	R1d
browse	browsed	browsed	d	aʊ	1	S1d
drown	drowned	drowned	d	aʊ	1	S1d
cure	cured	cured	d	ʊə	1	T2d
ruin	ruined	ruined	d	ʊɪ	1	U1d
shower	showered	showered	d	aʊə	1	V1d
fire	fired	fired	d	aɪə	1	W2d
lower	lowered	lowered	d	əʊə	1	Z1d
envy	envied	envied	d	e	2	A2d
measure	measured	measured	d	e	2	A2d
question	questioned	questioned	d	e	2	A2d
rebel	rebelled	rebelled	d	e	2	A2d
reckon	reckoned	reckoned	d	e	2	A2d
settle	settled	settled	d	e	2	A2d
schedule	scheduled	scheduled	d	e	2	A2d
strengthen	strengthened	strengthened	d	e	2	A2d
threaten	threatened	threatened	d	e	2	A2d
tremble	trembled	trembled	d	e	2	A2d
capture	captured	captured	d	æ	2	B2d
challenge	challenged	challenged	d	æ	2	B2d
gather	gathered	gathered	d	æ	2	B2d
tackle	tackled	tackled	d	æ	2	B2d

value	valued	valued	d	æ	2	B2d
double	doubled	doubled	d	ʌ	2	C2d
govern	governed	governed	d	ʌ	2	C2d
mumble	mumbled	mumbled	d	ʌ	2	C2d
mutter	muttered	muttered	d	ʌ	2	C2d
struggle	struggled	struggled	d	ʌ	2	C2d
suffer	suffered	suffered	d	ʌ	2	C2d
belong	belonged	belonged	d	ɒ	2	E2d
quarrel	quarrelled	quarrelled	d	ɒ	2	E2d
sponsor	sponsored	sponsored	d	ɒ	2	E2d
swallow	swallowed	swallowed	d	ɒ	2	E2d
wander	wandered	wandered	d	ɒ	2	E2d
differ	differed	differed	d	ɪ	2	G2d
shiver	shivered	shivered	d	ɪ	2	G2d
whisper	whispered	whispered	d	ɪ	2	G2d
whistle	whistled	whistled	d	ɪ	2	G2d
appeal	appealed	appealed	d	i:	2	H2d
deceive	deceived	deceived	d	i:	2	H2d
feature	featured	featured	d	i:	2	H2d
reveal	revealed	revealed	d	i:	2	H2d
burgle	burgled	burgled	d	ɜ:	2	I2d
concern	concerned	concerned	d	ɜ:	2	I2d
disturb	disturbed	disturbed	d	ɜ:	2	I2d
emerge	emerged	emerged	d	ɜ:	2	I2d
observe	observed	observed	d	ɜ:	2	I2d
occur	occurred	occurred	d	ɜ:	2	I2d
preserve	preserved	preserved	d	ɜ:	2	I2d
absorb	absorbed	absorbed	d	ɔ:	2	J2d
alter	altered	altered	d	ɔ:	2	J2d
ignore	ignored	ignored	d	ɔ:	2	J2d
recall	recalled	recalled	d	ɔ:	2	J2d
restore	restored	restored	d	ɔ:	2	J2d
transform	transformed	transformed	d	ɔ:	2	J2d
water	watered	watered	d	ɔ:	2	J2d
accuse	accused	accused	d	u:	2	K2d
amuse	amused	amused	d	u:	2	K2d
assume	assumed	assumed	d	u:	2	K2d
confuse	confused	confused	d	u:	2	K2d
consume	consumed	consumed	d	u:	2	K2d
Google	googled	googled	d	u:	2	K2d
renew	renewed	renewed	d	u:	2	K2d
review	reviewed	reviewed	d	u:	2	K2d
vary	varied	varied	d	eə	2	N2d
betray	betrayed	betrayed	d	eɪ	2	O2d
decay	decayed	decayed	d	eɪ	2	O2d
maintain	maintained	maintained	d	eɪ	2	O2d
obey	obeyed	obeyed	d	eɪ	2	O2d

obtain	obtained	obtained	d	eɪ	2	O2d
regain	regained	regained	d	eɪ	2	O2d
remain	remained	remained	d	eɪ	2	O2d
repay	repaid	repaid	d	eɪ	2	O2d
poison	poisoned	poisoned	d	ɔɪ	2	P2d
combine	combined	combined	d	aɪ	2	Q2d
decline	declined	declined	d	aɪ	2	Q2d
define	defined	defined	d	aɪ	2	Q2d
deny	denied	denied	d	aɪ	2	Q2d
deprive	deprived	deprived	d	aɪ	2	Q2d
disguise	disguised	disguised	d	aɪ	2	Q2d
frighten	frightened	frightened	d	aɪ	2	Q2d
resign	resigned	resigned	d	aɪ	2	Q2d
supply	supplied	supplied	d	aɪ	2	Q2d
survive	survived	survived	d	aɪ	2	Q2d
tighten	tightened	tightened	d	aɪ	2	Q2d
untie	untied	untied	d	aɪ	2	Q2d
widen	widened	widened	d	aɪ	2	Q2d
compose	composed	composed	d	əʊ	2	R2d
enclose	enclosed	enclosed	d	əʊ	2	R2d
expose	exposed	exposed	d	əʊ	2	R2d
oppose	opposed	opposed	d	əʊ	2	R2d
propose	proposed	proposed	d	əʊ	2	R2d
outline	outlined	outlined	d	aʊ	2	S2d
assure	assured	assured	d	ʊə	2	T2d
endure	endured	endured	d	ʊə	2	T2d
ensure	ensured	ensured	d	ʊə	2	T2d
acquire	acquired	acquired	d	aɪə	2	W2d
enquire	enquired	enquired	d	aɪə	2	W2d
inquire	inquired	inquired	d	aɪə	2	W2d
inspire	inspired	inspired	d	aɪə	2	W2d
emphasize	emphasized	emphasized	d	e	3	A3d
specialize	specialized	specialized	d	e	3	A3d
specify	specified	specified	d	e	3	A3d
terrify	terrified	terrified	d	e	3	A3d
abandon	abandoned	abandoned	d	æ	3	B3d
analyse	analysed	analysed	d	æ	3	B3d
examine	examined	examined	d	æ	3	B3d
satisfy	satisfied	satisfied	d	æ	3	B3d
discourage	discouraged	discouraged	d	ʌ	3	C3d
justify	justified	justified	d	ʌ	3	C3d
compromise	compromised	compromised	d	ɒ	3	E3d
occupy	occupied	occupied	d	ɒ	3	E3d
qualify	qualified	qualified	d	ɒ	3	E3d
criticize	criticized	criticized	d	ɪ	3	G3d
guarantee	guaranteed	guaranteed	d	i:	3	H3d
disapprove	disapproved	disapproved	d	u:	3	K3d

supervise	supervised	supervised	d	uː	3	K3d
interfere	interfered	interfered	d	ɪə	3	M3d
rearrange	rearranged	rearranged	d	eɪ	3	O2d
enable	enabled	enabled	d	eɪ	3	O3d
entitle	entitled	entitled	d	aɪ	3	Q3d
socialize	socialized	socialized	d	əʊ	3	R3d
encounter	encountered	encountered	d	aʊ	3	S3d
identify	identified	identified	d	e	4	A4d
manufacture	manufactured	manufactured	d	æ	4	B4d

Table 74. Similarity groups of English irregular verbs for the B2 level.

VERB	past tense	past participle	class	subcategory
bend	bent	bent	1	B
sweep	swept	swept	3	A
seek	sought	sought	3	B
bear	bore	borne	4	Ab
swear	swore	sworn	4	Ab
overtake	overtook	overtaken	4	Bb
forbid	forbad(e)	forbidden	4	Bc
rewrite	rewrote	rewritten	4	Ca
undo	undid	undone	4	Cc
broadcast	broadcast	broadcast	5	
burst	burst	burst	5	
split	split	split	5	
spread	spread	spread	5	
upset	upset	upset	5	
breed	bred	bred	6	A
misunderstand	misunderstood	misunderstood	6	H
shrink	shrank	shrunk	7	A
overcome	overcame	overcome	7	B

Table 75. Similarity groups of English regular verbs for the C1 level.

VERB	past tense	past participle	allomorph	vowel	syllables	class
cap	capped	capped	t	æ	1	B1t
rank	ranked	ranked	t	æ	1	B1t
tax	taxed	taxed	t	æ	1	B1t
dump	dumped	dumped	t	ʌ	1	C1t
hop	hopped	hopped	t	ɒ	1	E1t
stock	stocked	stocked	t	ɒ	1	E1t
swap	swapped	swapped	t	ɒ	1	E1t
sip	sipped	sipped	t	ɪ	1	G1t
group	grouped	grouped	t	uː	1	K1t

grasp	grasped	grasped	t	ɑː	1	L1t
march	marched	marched	t	ɑː	1	L1t
place	placed	placed	t	eɪ	1	O1t
trace	traced	traced	t	eɪ	1	O1t
price	priced	priced	t	aɪ	1	Q1t
address	addressed	addressed	t	e	2	A2t
possess	possessed	possessed	t	e	2	A2t
refresh	refreshed	refreshed	t	e	2	A2t
dismiss	dismissed	dismissed	t	ɪ	2	G2t
enrich	enriched	enriched	t	ɪ	2	G2t
rehearse	rehearsed	rehearsed	t	ɜː	2	I2t
worship	worshipped	worshipped	t	ɜː	2	I2t
enforce	enforced	enforced	t	ɔː	2	J2t
enhance	enhanced	enhanced	t	ɑː	2	L2t
displace	displaced	displaced	t	eɪ	2	O2t
embrace	embraced	embraced	t	eɪ	2	O2t
sacrifice	sacrificed	sacrificed	t	æ	3	B3t
accomplish	accomplished	accomplished	t	ɒ	3	E3t
diminish	diminished	diminished	t	ɪ	3	G3t
reinforce	reinforced	reinforced	t	ɔː	3	J3t
reproduce	reproduced	reproduced	t	uː	3	K3t
redevelop	redeveloped	redeveloped	t	e	4	A4t
bet	betted	betted	ɪd	e	1	A1ɪd
fund	funded	funded	ɪd	ʌ	1	C1ɪd
opt	opted	opted	ɪd	ɒ	1	E1ɪd
shift	shifted	shifted	ɪd	ɪ	1	G1ɪd
twist	twisted	twisted	ɪd	ɪ	1	G1ɪd
shield	shielded	shielded	ɪd	iː	1	H1ɪd
draft	drafted	drafted	ɪd	ɑː	1	L1ɪd
aid	aided	aided	ɪd	eɪ	1	O1ɪd
grade	graded	graded	ɪd	eɪ	1	O1ɪd
rate	rated	rated	ɪd	eɪ	1	O1ɪd
host	hosted	hosted	ɪd	əʊ	1	R1ɪd
quote	quoted	quoted	ɪd	əʊ	1	R1ɪd
crowd	crowded	crowded	ɪd	aʊ	1	S1ɪd
detect	detected	detected	ɪd	e	2	A2ɪd
digest	digested	digested	ɪd	e	2	A2ɪd
infect	infected	infected	ɪd	e	2	A2ɪd
inspect	inspected	inspected	ɪd	e	2	A2ɪd
neglect	neglected	neglected	ɪd	e	2	A2ɪd
impact	impacted	impacted	ɪd	æ	2	B2ɪd
consult	consulted	consulted	ɪd	ʌ	2	C2ɪd
corrupt	corrupted	corrupted	ɪd	ʌ	2	C2ɪd
instruct	instructed	instructed	ɪd	ʌ	2	C2ɪd
refund	refunded	refunded	ɪd	ʌ	2	C2ɪd
dictate	dictated	dictated	ɪd	ɪ	2	G2ɪd
omit	omitted	omitted	ɪd	ɪ	2	G2ɪd

restrict	restricted	restricted	ɪd	ɪ	2	G2ɪd
transmit	transmitted	transmitted	ɪd	ɪ	2	G2ɪd
exceed	exceeded	exceeded	ɪd	iː	2	H2ɪd
proceed	proceeded	proceeded	ɪd	iː	2	H2ɪd
assert	asserted	asserted	ɪd	ɜː	2	I2ɪd
insert	inserted	inserted	ɪd	ɜː	2	I2ɪd
applaud	applauded	applauded	ɪd	ɔː	2	J2ɪd
distort	distorted	distorted	ɪd	ɔː	2	J2ɪd
exhaust	exhausted	exhausted	ɪd	ɔː	2	J2ɪd
commute	commuted	commuted	ɪd	uː	2	K2ɪd
conclude	concluded	concluded	ɪd	uː	2	K2ɪd
exclude	excluded	excluded	ɪd	uː	2	K2ɪd
recruit	recruited	recruited	ɪd	uː	2	K2ɪd
restart	restarted	restarted	ɪd	ɑː	2	L2ɪd
appoint	appointed	appointed	ɪd	ɔɪ	2	P2ɪd
unite	united	United	ɪd	aɪ	2	Q2ɪd
unfold	unfolded	unfolded	ɪd	əʊ	2	R2ɪd
unload	unloaded	unloaded	ɪd	əʊ	2	R2ɪd
recreate	recreated	recreated	ɪd	e	3	A3ɪd
regulate	regulated	regulated	ɪd	e	3	A3ɪd
renovate	renovated	renovated	ɪd	e	3	A3ɪd
allocate	allocated	allocated	ɪd	æ	3	B3ɪd
fascinate	fascinated	fascinated	ɪd	æ	3	B3ɪd
cultivate	cultivated	cultivated	ɪd	ʌ	3	C3ɪd
readjust	readjusted	readjusted	ɪd	ʌ	3	C3ɪd
reconstruct	reconstructed	reconstructed	ɪd	ʌ	3	C3ɪd
compensate	compensated	compensated	ɪd	ɒ	3	E3ɪd
complicate	complicated	complicated	ɪd	ɒ	3	E3ɪd
compliment	complimented	complimented	ɪd	ɒ	3	E3ɪd
constitute	constituted	constituted	ɪd	ɒ	3	E3ɪd
deposit	deposited	deposited	ɪd	ɒ	3	E3ɪd
nominate	nominated	nominated	ɪd	ɒ	3	E3ɪd
contradict	contradicted	contradicted	ɪd	ɪ	3	G3ɪd
exhibit	exhibited	exhibited	ɪd	ɪ	3	G3ɪd
imitate	imitated	imitated	ɪd	ɪ	3	G3ɪd
integrate	integrated	integrated	ɪd	ɪ	3	G3ɪd
irritate	irritated	irritated	ɪd	ɪ	3	G3ɪd
alternate	alternated	alternated	ɪd	ɜː	3	I3ɪd
relocate	relocated	relocated	ɪd	eɪ	3	O3ɪd
motivate	motivated	motivated	ɪd	əʊ	3	R3ɪd
accelerate	accelerated	accelerated	ɪd	e	4	A4ɪd
collaborate	collaborated	collaborated	ɪd	æ	4	B4ɪd
evaluate	evaluated	evaluated	ɪd	æ	4	B4ɪd
exaggerate	exaggerated	exaggerated	ɪd	æ	4	B4ɪd
anticipate	anticipated	anticipated	ɪd	ɪ	4	G4ɪd
discriminate	discriminated	discriminated	ɪd	ɪ	4	G4ɪd
eliminate	eliminated	eliminated	ɪd	ɪ	4	G4ɪd

facilitate	facilitated	facilitated	ɪd	ɪ	4	G4ɪd
accumulate	accumulated	accumulated	ɪd	uː	4	K4ɪd
deteriorate	deteriorated	deteriorated	ɪd	ɪə	4	M4ɪd
associate	associated	associated	ɪd	əʊ	4	R4ɪd
negotiate	negotiated	negotiated	ɪd	əʊ	4	R4ɪd
differentiate	differentiated	differentiated	ɪd	e	5	A5ɪd
overestimate	overestimated	overestimated	ɪd	e	5	A5ɪd
jam	jammed	jammed	d	æ	1	B1d
scan	scanned	scanned	d	æ	1	B1d
plunge	plunged	plunged	d	ʌ	1	C1d
surge	surged	surged	d	ɜː	1	I1d
sue	sued	sued	d	uː	1	K1d
starve	starved	starved	d	ɑː	1	L1d
rear	reared	reared	d	ɪə	1	M1d
scare	scared	scared	d	eə	1	N1d
spare	spared	spared	d	eə	1	N1d
bribe	bribed	bribed	d	aɪ	1	Q1d
spy	spied	spied	d	aɪ	1	Q1d
thrive	thrived	thrived	d	aɪ	1	Q1d
pose	posed	posed	d	əʊ	1	R1d
sow	sowed	sowed	d	əʊ	1	R1d
lengthen	lengthened	lengthened	d	e	2	A2d
lessen	lessened	lessened	d	e	2	A2d
render	rendered	rendered	d	e	2	A2d
narrow	narrowed	narrowed	d	æ	2	B2d
puzzle	puzzled	puzzled	d	ʌ	2	C2d
smuggle	smuggled	smuggled	d	ʌ	2	C2d
number	numbered	numbered	d	ʌ	2	C2d
bully	bullied	bullied	d	ʊ	2	D2d
conquer	conquered	conquered	d	ɒ	2	E2d
dissolve	dissolved	dissolved	d	ɒ	2	E2d
evolve	evolved	evolved	d	ɒ	2	E2d
prolong	prolonged	prolonged	d	ɒ	2	E2d
resolve	resolved	resolved	d	ɒ	2	E2d
mingle	mingled	mingled	d	ɪ	2	G2d
picture	pictured	pictured	d	ɪ	2	G2d
trigger	triggered	triggered	d	ɪ	2	G2d
deepen	deepened	deepened	d	iː	2	H2d
perceive	perceived	perceived	d	iː	2	H2d
weaken	weakened	weakened	d	iː	2	H2d
worsen	worsened	worsened	d	ɜː	2	I2d
broaden	broadened	broadened	d	ɔː	2	J2d
shorten	shortened	shortened	d	ɔː	2	J2d
misuse	misused	misused	d	uː	2	K2d
presume	presumed	presumed	d	uː	2	K2d
pursue	pursued	pursued	d	uː	2	K2d
resume	resumed	resumed	d	uː	2	K2d

alarm	alarmed	alarmed	d	ɑː	2	L2d
recharge	recharged	recharged	d	ɑː	2	L2d
beware	bewared	bewared	d	eə	2	N2d
attain	attained	attained	d	eɪ	2	O2d
campaign	campaigned	campaigned	d	eɪ	2	O2d
cater	catered	catered	d	eɪ	2	O2d
convey	conveyed	conveyed	d	eɪ	2	O2d
daydream	daydreamed	daydreamed	d	eɪ	2	O2d
engage	engaged	engaged	d	eɪ	2	O2d
label	labelled	labelled	d	eɪ	2	O2d
outweigh	outweighed	outweighed	d	eɪ	2	O2d
restrain	restrained	restrained	d	eɪ	2	O2d
survey	surveyed	surveyed	d	eɪ	2	O2d
assign	assigned	assigned	d	aɪ	2	Q2d
compile	compiled	compiled	d	aɪ	2	Q2d
comply	complied	complied	d	aɪ	2	Q2d
comprise	comprised	comprised	d	aɪ	2	Q2d
oblige	obliged	obliged	d	aɪ	2	Q2d
impose	imposed	imposed	d	əʊ	2	R2d
suppose	supposed	supposed	d	əʊ	2	R2d
total	totalled	totalled	d	əʊ	2	R2d
outrage	outraged	outraged	d	aʊ	2	S2d
desire	desired	desired	d	aɪə	2	W2d
overwhelm	overwhelmed	overwhelmed	d	e	3	A3d
rectify	rectified	rectified	d	e	3	A3d
resemble	resembled	resembled	d	e	3	A3d
verify	verified	verified	d	e	3	A3d
clarify	clarified	clarified	d	æ	3	B3d
classify	classified	classified	d	æ	3	B3d
outnumber	outnumbered	outnumbered	d	ʌ	3	C3d
summarize	summarized	summarized	d	ʌ	3	C3d
uncover	uncovered	uncovered	d	ʌ	3	C3d
acknowledge	acknowledged	acknowledged	d	ɒ	3	E3d
modify	modified	modified	d	ɒ	3	E3d
monitor	monitored	monitored	d	ɒ	3	E3d
imprison	imprisoned	imprisoned	d	ɪ	3	G2d
envisage	envisaged	envisaged	d	ɪ	3	G3d
envision	envisioned	envisioned	d	ɪ	3	G3d
minimize	minimized	minimized	d	ɪ	3	G3d
simplify	simplified	simplified	d	ɪ	3	G3d
determine	determined	determined	d	ɜː	3	I3d
authorize	authorized	authorized	d	ɔː	3	J3d
misinform	misinformed	misinformed	d	ɔː	3	J3d
volunteer	volunteered	volunteered	d	ɪə	3	M3d
misbehave	misbehaved	misbehaved	d	eɪ	3	O3d
privatize	privatized	privatized	d	aɪ	3	Q3d

notify	notified	notified	d	əʊ	3	R3d
reassure	reassured	reassured	d	ʊə	3	T3d
generalize	generalized	generalized	d	e	4	A4d
reconsider	reconsidered	reconsidered	d	ɪ	4	G4d
reorganize	reorganized	reorganized	d	ɔː	4	J4d

Table 76. Similarity groups of English irregular verbs for the C1 level.

VERB	R/IR	past tense	past participle	class	subcategory
sow	IR	sowed	sown	2	
daydream	IR	daydreamt	daydreamt	3	A
flee	IR	fled	fled	3	A
rethink	IR	rethought	rethought	3	B
undertake	IR	undertook	undertaken	4	Bb
withdraw	IR	withdrew	withdrawn	4	Bd
foresee	IR	foresaw	foreseen	4	Bg
arise	IR	arose	arisen	4	Ca
overdo	IR	overdid	overdone	4	Cc
bet	IR	bet	bet	5	
mislead	IR	misled	misled	6	A
spin	IR	span/spun	spun	6	B
unwind	IR	unwound	unwound	6	C
undergo	IR	underwent	undergone	7	C

Table 77. Similarity groups of English regular verbs for the C2 level.

VERB	past tense	past participle	allomorph	vowel	syllables	class
sense	sensed	sensed	t	e	1	A1t
wreck	wrecked	wrecked	t	e	1	A1t
back	backed	backed	t	æ	1	B1t
clash	clashed	clashed	t	æ	1	B1t
patch	patched	patched	t	æ	1	B1t
scrap	scrapped	scrapped	t	æ	1	B1t
snap	snapped	snapped	t	æ	1	B1t
snatch	snatched	snatched	t	æ	1	B1t
track	tracked	tracked	t	æ	1	B1t
clutch	clutched	clutched	t	ʌ	1	C1t
crush	crushed	crushed	t	ʌ	1	C1t
suck	sucked	sucked	t	ʌ	1	C1t
pop	popped	popped	t	ɒ	1	E1t
rock	rocked	rocked	t	ɒ	1	E1t
top	topped	topped	t	ɒ	1	E1t
sniff	sniffed	sniffed	t	ɪ	1	G1t
tip	tipped	tipped	t	ɪ	1	G1t

wink	winked	winked	t	ɪ	1	G1t
squeak	squeaked	squeaked	t	iː	1	H1t
nurse	nursed	nursed	t	ɜː	1	I1t
arch	arched	arched	t	ɑː	1	L1t
gasp	gasped	gasped	t	ɑː	1	L1t
spark	sparked	sparked	t	ɑː	1	L1t
fake	faked	faked	t	eɪ	1	O1t
grace	graced	graced	t	eɪ	1	O1t
shape	shaped	shaped	t	eɪ	1	O1t
voice	voiced	voiced	t	ɔɪ	1	P1t
dice	diced	diced	t	aɪ	1	Q1t
commence	commenced	commenced	t	e	2	A2t
relish	relished	relished	t	e	2	A2t
suppress	suppressed	suppressed	t	e	2	A2t
elapse	elapsed	elapsed	t	æ	2	B2t
flourish	flourished	flourished	t	ʌ	2	C2t
endorse	endorsed	endorsed	t	ɔː	2	J2t
deduce	deduced	deduced	t	uː	2	K2t
seduce	seduced	seduced	t	uː	2	K2t
surpass	surpassed	surpassed	t	ɑː	2	L2t
misplace	misplaced	misplaced	t	eɪ	2	O2t
provoke	provoked	provoked	t	əʊ	2	R2t
reproach	reproached	reproached	t	əʊ	2	R2t
embarrass	embarrassed	embarrassed	t	æ	3	B3t
overlap	overlapped	overlapped	t	æ	3	B3t
dread	dreaded	dreaded	ɪd	e	1	A1ɪd
nest	nested	nested	ɪd	e	1	A1ɪd
tread	treaded	treaded	ɪd	e	1	A1ɪd
pat	patted	patted	ɪd	æ	1	B1ɪd
bond	bonded	bonded	ɪd	ɒ	1	E1ɪd
rot	rotted	rotted	ɪd	ɒ	1	E1ɪd
drift	drifted	drifted	ɪd	ɪ	1	G1ɪd
plead	pleaded	pleaded	ɪd	iː	1	H1ɪd
wield	wielded	wielded	ɪd	iː	1	H1ɪd
yield	yielded	yielded	ɪd	iː	1	H1ɪd
part	parted	parted	ɪd	ɑː	1	L1ɪd
raid	raided	raided	ɪd	eɪ	1	O1ɪd
cloud	clouded	clouded	ɪd	aʊ	1	S1ɪd
mount	mounted	mounted	ɪd	aʊ	1	S1ɪd
amend	amended	amended	ɪd	e	2	A2ɪd
consent	consented	consented	ɪd	e	2	A2ɪd
merit	merited	merited	ɪd	e	2	A2ɪd
perfect	perfected	perfected	ɪd	e	2	A2ɪd
resent	resented	resented	ɪd	e	2	A2ɪd
contract	contracted	contracted	ɪd	æ	2	B2ɪd
comfort	comforted	comforted	ɪd	ʌ	2	C2ɪd
confront	confronted	confronted	ɪd	ʌ	2	C2ɪd

erupt	erupted	erupted	ɪd	ʌ	2	C2ɪd
combat	combated	combated	ɪd	ɒ	2	E2ɪd
vomit	vomited	vomited	ɪd	ɒ	2	E2ɪd
convict	convicted	convicted	ɪd	ɪ	2	G2ɪd
depict	depicted	depicted	ɪd	ɪ	2	G2ɪd
emit	emitted	emitted	ɪd	ɪ	2	G2ɪd
persist	persisted	persisted	ɪd	ɪ	2	G2ɪd
pinpoint	pinpointed	pinpointed	ɪd	ɪ	2	G2ɪd
concede	conceded	conceded	ɪd	iː	2	H2ɪd
precede	preceded	preceded	ɪd	iː	2	H2ɪd
retreat	retreated	retreated	ɪd	iː	2	H2ɪd
alert	alerted	alerted	ɪd	ɜː	2	I2ɪd
divert	diverted	diverted	ɪd	ɜː	2	I2ɪd
assault	assaulted	assaulted	ɪd	ɔː	2	J2ɪd
dispute	disputed	disputed	ɪd	uː	2	K2ɪd
intrude	intruded	intruded	ɪd	uː	2	K2ɪd
target	targeted	targeted	ɪd	ɑː	2	L2ɪd
contrast	contrasted	contrasted	ɪd	ɑː/æ	2	L2ɪd/B2ɪd
debate	debated	debated	ɪd	eɪ	2	O2ɪd
equate	equated	equated	ɪd	eɪ	2	O2ɪd
relate	related	related	ɪd	eɪ	2	O2ɪd
riot	rioted	rioted	ɪd	aɪ	2	Q2ɪd
erode	eroded	eroded	ɪd	əʊ	2	R2ɪd
revolt	revolted	revolted	ɪd	əʊ	2	R2ɪd
delegate	delegated	delegated	ɪd	e	3	A3ɪd
execute	executed	executed	ɪd	e	3	A3ɪd
inherit	inherited	inherited	ɪd	e	3	A3ɪd
recollect	recollected	recollected	ɪd	e	3	A3ɪd
speculate	speculated	speculated	ɪd	e	3	A3ɪd
advocate	advocated	advocated	ɪd	æ	3	B3ɪd
inhabit	inhabited	inhabited	ɪd	æ	3	B3ɪd
vaccinate	vaccinated	vaccinated	ɪd	æ	3	B3ɪd
fluctuate	fluctuated	fluctuated	ɪd	ʌ	3	C3ɪd
supplement	supplemented	supplemented	ɪd	ʌ	3	C3ɪd
contemplate	contemplated	contemplated	ɪd	ɒ	3	E3ɪd
moderate	moderated	moderated	ɪd	ɒ	3	E3ɪd
prosecute	prosecuted	prosecuted	ɪd	ɒ	3	E3ɪd
formulate	formulated	formulated	ɪd	ɔː	3	J3ɪd
coincide	coincided	coincided	ɪd	aɪ	3	Q3ɪd
violate	violated	violated	ɪd	aɪ	3	Q3ɪd
overload	overloaded	overloaded	ɪd	əʊ	3	R3ɪd
commemorate	commemorated	commemorated	ɪd	e	4	A4ɪd
eradicate	eradicated	eradicated	ɪd	æ	4	B4ɪd
initiate	initiated	initiated	ɪd	ɪ	4	G4ɪd
redistribute	redistributed	redistributed	ɪd	ɪ	4	G4ɪd
misinterpret	misinterpreted	misinterpreted	ɪd	ɜː	4	I4ɪd
incorporate	incorporated	incorporated	ɪd	ɔː	4	J4ɪd

buzz	buzzed	buzzed	d	ʌ	1	C1d
hum	hummed	hummed	d	ʌ	1	C1d
shrug	shrugged	shrugged	d	ʌ	1	C1d
dim	dimmed	dimmed	d	ɪ	1	G1d
grin	grinned	grinned	d	ɪ	1	G1d
thrill	thrilled	thrilled	d	ɪ	1	G1d
deem	deemed	deemed	d	iː	1	H1d
seal	sealed	sealed	d	iː	1	H1d
curb	curbed	curbed	d	ɜː	1	I1d
merge	merged	merged	d	ɜː	1	I1d
spur	spurred	spurred	d	ɜː	1	I1d
urge	urged	urged	d	ɜː	1	I1d
haul	hauled	hauled	d	ɔː	1	J1d
roar	roared	roared	d	ɔː	1	J1d
soar	soared	soared	d	ɔː	1	J1d
glue	glued	glued	d	uː	1	K1d
loom	loomed	loomed	d	uː	1	K1d
view	viewed	viewed	d	uː	1	K1d
carve	carved	carved	d	ɑː	1	L1d
charm	charmed	charmed	d	ɑː	1	L1d
scar	scarred	scarred	d	ɑː	1	L1d
glare	glared	glared	d	eə	1	N1d
crave	craved	craved	d	eɪ	1	O1d
drain	drained	drained	d	eɪ	1	O1d
hail	hailed	hailed	d	eɪ	1	O1d
plague	plagued	plagued	d	eɪ	1	O1d
rage	raged	raged	d	eɪ	1	O1d
reign	reigned	reigned	d	eɪ	1	O1d
stain	stained	stained	d	eɪ	1	O1d
trail	trailed	trailed	d	eɪ	1	O1d
coin	coined	coined	d	ɔɪ	1	P1d
line	lined	lined	d	aɪ	1	Q1d
rhyme	rhymed	rhymed	d	aɪ	1	Q1d
strive	strived	strived	d	aɪ	1	Q1d
glow	glowed	glowed	d	əʊ	1	R1d
groan	groaned	groaned	d	əʊ	1	R1d
moan	moaned	moaned	d	əʊ	1	R1d
slow	slowed	slowed	d	əʊ	1	R1d
loathe	loathed	loathed	ɪd	əʊ	1	R1d
frown	frowned	frowned	d	aʊ	1	S1d
house	housed	housed	d	aʊ	1	S1d
vow	vowed	vowed	d	aʊ	1	S1d
lure	lured	lured	d	ʊə	1	T1d
tire	tired	tired	d	aɪə	1	W1d
allege	alleged	alleged	d	e	2	A2d
condemn	condemned	condemned	d	e	2	A2d
echo	echoed	echoed	d	e	2	A2d

excel	excelled	excelled	d	e	2	A2d
gesture	gestured	gestured	d	e	2	A2d
lecture	lectured	lectured	d	e	2	A2d
treasure	treasured	treasured	d	e	2	A2d
battle	battled	battled	d	æ	2	B2d
anchor	anchored	anchored	d	æ	2	B2d
blackmail	blackmailed	blackmailed	d	æ	2	B2d
channel	channelled	channelled	d	æ	2	B2d
dazzle	dazzled	dazzled	d	æ	2	B2d
flatten	flattened	flattened	d	æ	2	B2d
gamble	gambled	gambled	d	æ	2	B2d
sadden	saddened	saddened	d	æ	2	B2d
saddle	saddled	saddled	d	æ	2	B2d
stagger	staggered	staggered	d	æ	2	B2d
bubble	bubbled	bubbled	d	ʌ	2	C2d
crumble	crumbled	crumbled	d	ʌ	2	C2d
indulge	indulged	indulged	d	ʌ	2	C2d
juggle	juggled	juggled	d	ʌ	2	C2d
shudder	shuddered	shuddered	d	ʌ	2	C2d
structure	structured	structured	d	ʌ	2	C2d
stumble	stumbled	stumbled	d	ʌ	2	C2d
succumb	succumbed	succumbed	d	ʌ	2	C2d
summon	summoned	summoned	d	ʌ	2	C2d
utter	uttered	uttered	d	ʌ	2	C2d
lobby	lobbied	lobbied	d	ɒ	2	E2d
ponder	pondered	pondered	d	ɒ	2	E2d
prosper	prospered	prospered	d	ɒ	2	E2d
squander	squandered	squandered	d	ɒ	2	E2d
figure	figured	figured	d	ɪ	2	G2d
filter	filtered	filtered	d	ɪ	2	G2d
giggle	giggled	giggled	d	ɪ	2	G2d
hinder	hindered	hindered	d	ɪ	2	G2d
issue	issued	issued	d	ɪ	2	G2d
linger	lingered	lingered	d	ɪ	2	G2d
pity	pitied	pitied	d	ɪ	2	G2d
signal	signalled	signalled	d	ɪ	2	G2d
sprinkle	sprinkled	sprinkled	d	ɪ	2	G2d
conceal	concealed	concealed	d	iː	2	H2d
conceive	conceived	conceived	d	iː	2	H2d
relieve	relieved	relieved	d	iː	2	H2d
retrieve	retrieved	retrieved	d	iː	2	H2d
further	furthered	furthered	d	ɜː	2	I2d
incur	incurred	incurred	d	ɜː	2	I2d
infer	inferred	inferred	d	ɜː	2	I2d
murmur	murmured	murmured	d	ɜː	2	I2d
recur	recurred	recurred	d	ɜː	2	I2d
caution	cautioned	cautioned	d	ɔː	2	J2d

reform	reformed	reformed	d	ɔː	2	J2d
torture	tortured	tortured	d	ɔː	2	J2d
loosen	loosened	loosened	d	uː	2	K2d
sample	sampled	sampled	d	ɑː	2	L2d
sharpen	sharpened	sharpened	d	ɑː	2	L2d
constrain	constrained	constrained	d	eɪ	2	O2d
exclaim	exclaimed	exclaimed	d	eɪ	2	O2d
favour	favoured	favoured	d	eɪ	2	O2d
portray	portrayed	portrayed	d	eɪ	2	O2d
prevail	prevailed	prevailed	d	eɪ	2	O2d
refrain	refrained	refrained	d	eɪ	2	O2d
retain	retained	retained	d	eɪ	2	O2d
sustain	sustained	sustained	d	eɪ	2	O2d
confine	confined	confined	d	aɪ	2	Q2d
defy	defied	defied	d	aɪ	2	Q2d
devise	devised	devised	d	aɪ	2	Q2d
imply	implied	implied	d	aɪ	2	Q2d
prescribe	prescribed	prescribed	d	aɪ	2	Q2d
refine	refined	refined	d	aɪ	2	Q2d
revive	revived	revived	d	aɪ	2	Q2d
rival	rivalled	rivalled	d	aɪ	2	Q2d
disclose	disclosed	disclosed	d	əʊ	2	R2d
arouse	aroused	aroused	d	aʊ	2	S2d
expire	expired	expired	d	aɪə	2	W2d
assemble	assembled	assembled	d	e	3	A3d
reconcile	reconciled	reconciled	d	e	3	A3d
surrender	surrendered	surrendered	d	e	3	A3d
testify	testified	testified	d	e	3	A3d
maximize	maximized	maximized	d	æ	3	B3d
restructure	restructured	restructured	d	ʌ	3	C3d
embody	embodied	embodied	d	ɒ	3	E3d
condition	conditioned	conditioned	d	ɪ	3	G3d
sympathize	sympathized	sympathized	d	ɪ	3	G3d
intervene	intervened	intervened	d	iː	3	H3d
diagnose	diagnosed	diagnosed	d	aɪ	3	Q3d
undermine	undermined	undermined	d	aɪ	3	Q3d
overflow	overflowed	overflowed	d	əʊ	3	R3d
exemplify	exemplified	exemplified	d	e	4	A4d
intensify	intensified	intensified	d	e	4	A4d
disqualify	disqualified	disqualified	d	ɒ	4	E4d

Table 78. Similarity groups of English irregular verbs for the C2 level.

VERB	past tense	past participle	class	subcategory
weep	wept	wept	3	A
overhear	overheard	overheard	3	E
bid	bid	bid	5	
cast	cast	cast	5	
offset	offset	offset	5	
uphold	upheld	upheld	6	A
cling	clung	clung	6	B
bind	bound	bound	6	C
spit	spat/spit	spat/spit	6	E
withstand	withstood	withstood	6	H