Filozofická fakulta Univerzity Palackého

Phonetic category formation in child EFL learners: development and organisation of L2 vowels

(Bakalářská práce)

2023

Monika Kučerová

Filozofická fakulta Univerzity Palackého
Katedra anglistiky a amerikanistiky

Phonetic category formation in child EFL learners: development and organisation of L2 vowels (Bakalářská práce).

Autor: Monika Kučerová
Studijní obor: Anglická filologie
Vedoucí práce: Mgr. Šárka Šimáčková, Ph.D.
Počet znaků: 84816
Počet stran: 47
Olomouc 2023

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a uvedla úplný seznam citované a použité literatury.

V Olomouci dne 5. 5. 2023 Monika Kučerová

I would like to express my gratitude to my supervisor Mgr. Šárka Šimáčková, Ph.D., and to Mgr. Václav Jonáš Podlipský, Ph.D. for their guidance, support, and encouragement throughout the writing of this thesis. I am also grateful to all the children who participated in this study and their parents.

Abstract

High quality input during L2 learning is considered crucial for L2 category development, especially in early stages of learning. Non-native speaker models, highly prevalent in EFL contexts, may adversely affect L2 category development. This study investigated the impact of input quality on production of L2 vowels /i, ɪ, ɛ, æ, ʌ/ by child learners. They produced words in a picture-naming task on six occasions throughout three months. The vowels /ɛ, æ/ were produced as better separated in quality in words that were heard primarily in input that preserved the /æ-ɛ/ contrast than in words from input that did not preserve the contrast. Despite /æ, ɛ, ʌ/ being produced as different in words heard in lessons, /æ/ still showed a large overlap with both /ɛ, ʌ/, indicating an underdeveloped category used in the L2. Input source did not affect the production of /i, ɪ/. The vowel /i/ was the only one that improved significantly with time, though it did not contribute to a better separation of /i, ɪ/ in quality. This study further observed change in the production of L1 vowels during a 10-week period, suggesting phonetic drift in inexperienced child EFL learners. Additionally, a task focused at judging the acceptability of word mispronunciations with substituted L2 vowels was conducted, with the results showing great individual variability in the learners' performance. Some accepted most pronunciations, while others accepted most words and rejected most nonwords, with similar scores for all vowels tested.

Keywords

vowel production, L2 contrasts, phonetic drift, foreign language acquisition, child L2 learning, non-native input

List of abbreviations

EFL	English as a Foreign Language
ERB	Equivalent rectangular bandwidth
F0	fundamental frequency
F1	first formant
F2	second formant
FLA	Foreign Language
FL	Foreign Language
L1	First Language
L2	Second Language
SLA	Second Language Acquisition
SSBE	Standard Southern British English

Table of contents

1 Introduction

Offering to provide children with a head start in language learning, foreign language courses for pre-schoolers have become a trend in recent years. However, research on L2 vowel acquisition by young children learning EFL is sparse. This study examines vowel production and perception of preschoolers attending EFL lessons, focusing on the L2 vowels /i, ɪ, ɛ, æ, ʌ/ elicited in single-word productions in six sessions across three months. By exploring the L2 vowel production and perception of inexperienced child learners, this study aims to account for how foreign accented L2 input that differs in the degree of preservation of L2 vowel contrasts impacts the learners' L2 vowel production, and how L1 vowels change during L2 learning. Foreign accented input that does not conserve the differences between distinct L2 sounds may inhibit the learners' L2 category development, and in the case of distinctive vowel categories, the learners' pronunciation could lead to decreased intelligibility. Investigating speech production of L2 learners in an L1-dominant environment allows for assessing the difficulties these child learners encounter and can provide a basis for teaching EFL in a way that helps children overcome the constraints of the environment. Additionally, the study investigates whether limited L2 learning can trigger changes in production of L1 vowels, focusing on the Moravian Czech /iː, i, ɛ, a/ recorded in single-word productions on two occasions separated by 10 weeks.

The thesis uses The Perceptual Assimilation Model of L2 speech learning (Best and Tyler 2007) to formulate predictions regarding the children's L2 vowel production and perception, referring to vowels in their L1. The children are exposed to two sources of input, one of which does not preserve a relevant vowel contrast, and the other which does preserve it. The learners produced words they had learned months ago and practiced at home, as well as words learned 0-3 with the teacher weeks before the recording. To account for the duality of input, words produced by the children were divided according to the input, which is expected to have provided most instances of the words. The learners' differentiation of the vowels in production was predicted based on Pierrehumbert's (2001) exemplar-based model, which presumes a connection of the learners' lexical and phonetic representations, allowing predictions to be formed separately for words known from either source of input. Lastly, research on phonetic drift in adults and children with little L2 experience is reviewed and related to the learners in this study.

The learners examined in this study were children who had been attending 45-minute weekly EFL lessons for at least 5 months at the start of the experiment. Though some had been learning English for more than 2 years, all were low-proficiency learners attending beginner lessons. None of the children was exposed to native speaker English interactively. All were taught by a proficient Czech-English bilingual, and practised English at home.

2 Literature review

2.1 SLM-r and PAM-L2

This study focuses on pre-schoolers acquiring EFL in a non-immersive L1-dominant environment. Predictions of L2 vowel development are based on the Perceptual Assimilation Model of L2 speech learning (PAM-L2, Best and Tyler 2007), and complemented with information from the Speech Learning Model and its revised version (SLM, Flege 1995; SLM-r, Flege and Bohn 2021). Although both deal with SLA, they have been used in FLA contexts (Šimáčková and Podlipský 2018, Šturm and Skarnitzl 2011). Still, even though they have been successfully used to predict L2 development in non-immersive settings, FLA as an impoverished context for L2 learning, should not be conflated with SLA findings. SLA typically occurs in natural communicative contexts where the target language is dominant, while FLA takes place in settings where the target language is not widely used and is commonly taught by non-native speakers or speakers from diverse L2 varieties in a classroom setting, possibly resulting in variable or inaccurate models of L2 pronunciation for FLA learners (Flege and Eefting 1987). This study focuses on FL learner's acquisition of the English vowels /i, ɪ, ɛ, æ, ʌ/.

Both models are based on the central idea that L1 and L2 sound categories share a "common phonological space" (Flege 1995, 239). SLM(-r) aims to account for mechanisms influencing both production and perception of L2 sounds, focusing on the reorganisation of individuals' phonetic systems as a function of the "phonetic input received" during L2 learning (Flege and Bohn 2021, 64). SLM-r posits that the acquisition of a new L2 category depends on the perceived proximity of the L2 sounds in the input to the closest L1 category: L2 sounds that are perceived as different from the L1 category are more probable to be attained than those that are perceived as resembling the L1 category (Flege and Bohn 2021). As the learners in this study are inexperienced, their L1 is expected to exert a strong influence on their L2. They are also learning EFL, which makes their chances for new L2 category formation slim (Tyler 2019). In their case, aiming to distinguish contrasting L2 vowels using distinctive features relevant to the L2 is an appropriate goal, even if the realisations of individual vowels are non-native-like. PAM-L2 aims to predict discrimination difficulty levels for L2 contrasts based on cross-language assimilation patterns (Best and Tyler 2007). To formulate predictions for the investigated L2 contrasts (/i-ɪ/, /ɛ-æ/, /æ-ʌ/) regarding their discrimination and preservation in production, it is helpful to examine the vowels in the learners' L1, and determine the pattern in which they map onto the L2 vowels in their input.

2.2 Comparison of non-back vowels in the learners' L1 and L2

This study focuses on learners acquiring Standard Southern British English (SSBE). Cruttenden (2014) describes SSBE as having 12 monophthongs, out of which five are relevant for this study: the front vowels /i, ɪ, ɛ, æ/ and the mid-low central vowel /ʌ/. SSBE vowels typically differ in length, which is, however, not a distinctive feature, hence the vowels are described as lax or tense, rather than long or short. Among the five vowels of interest, four are classified as lax (/ɪ, ɛ, æ, ʌ/), and /i/ as tense. The vowel /æ/ is longer than other lax vowels in SSBE, especially in open syllables and before voiced consonants.

Recently, /æ/ has become more open in SSBE, moving closer to /ʌ/. In the present study, this tendency is reflected in the teacher's pronunciation.

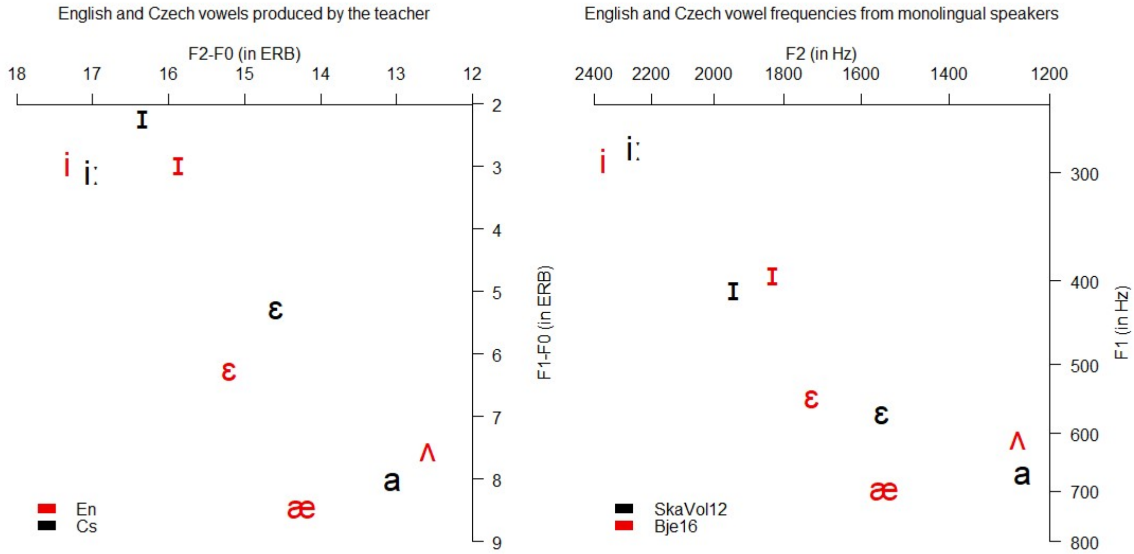
In Czech, quantity is a contrastive feature, and its five vowel qualities occur in two degrees of duration (Skarnitzl, Šturm and Volín 2016). In the non-back portion of the vowel space, Moravian Czech has only three vowels differentiated primarily by quality (/i, ɛ, a/) (Šimáčková, Podlipský and Chládková 2012), while English has five (/i, ɪ, ɛ, æ, ʌ/). Complex vowel systems like the English one make the acquisition of English vowels difficult for learners whose L1 vowel system is simpler. This is because a given pair of adjacent English vowels that differ qualitatively will likely map onto a single Czech vowel quality, with learners reusing the quantity cue to distinguish the L2 vowels (Šimáčková and Podlipský 2018).

As cue weighing is language-specific, and so must be learned by the L2 learner, the use of L1 cues in the L2, at least at the early stages of L2 learning, is virtually unavoidable (Morrison 2009). This study does not investigate the learners' use of vowel duration, but their potential use of it as a primary cue in distinguishing some L2 vowel contrasts may influence whether those vowels are differentiated in quality. Native English speakers rely on vowel quality and quantity when categorising English vowels (Escudero 2001), but as duration is not a contrastive feature, temporal cues can be absent on reduced in speech.

This study investigates speakers whose L1 is Moravian Czech. Focusing on quality, the vowels that are most similar across the languages are /iː/ and /i/ (Figure 1). For the mid-high front vowel, the symbol /ɪ/ is typically used in Bohemian Czech, where speakers produce the vowel as more separated in quality from /i/ than speakers of Moravian Czech (Šimáčková, Podlipský and Chládková 2012). Bohemian speakers also primarily rely on the quality cue in perception, with quantity only as a secondary cue, unlike Moravian speakers, who rely primarily on the quantity cue, with the quality cue being secondary (Podlipský, Skarnitzl and Volín 2009). The symbol /ɪ/ is used here for the Moravian Czech vowel typically transcribed as /i/ (Šimáčková, Podlipský and Chládková 2012) to highlight the expected mapping pattern and avoid confusion with the English /i/ and Czech /iː/, although the Moravian /ɪ~i/ is expected to map onto the English /ɪ/ based on duration, not quality. The Moravian /i/ is realised as higher and more front than the English /ɪ/. Also to highlight the mapping pattern, the symbol for the Moravian Czech mid front vowel used here is /ɛ/, though it is realised as more closed in Moravian than in English, and so is sometimes transcribed as /e/. The SSBE /ʌ/ is close to the Moravian /a/ in quality, though /a/ is lower. The SSBE /æ/ is also similar to /a/, but noticeably more front.

Figure 1 shows vowels produced by the learners' teacher, a speaker of Moravian Czech, for lack of data on monolingual Moravian speakers, and also to illustrate the vowels the learners were exposed to. The right plot is used here to show how the vowels are produced by native speakers; it must be noted, however, that the learners were not exposed to native speaker input. The teacher preserved the qualitative differences between English vowels, though especially in the case of /ɪ/, she did not do so in a native-like way.

Figure 1. **Left**: Czech and English vowels produced by the teacher (L1 Moravian Czech). **Right**: F1 and F2 for Czech (reported in Skarnitzl and Volín 2012) and SSBE (reported in Bjelaković 2016) non-back vowels. Axes are scaled logarithmically.



It is important to refer specifically to Moravian Czech vowels, as differences in L2-L1 mapping arise due to differences in L1 category organization, which can be attributed to exposure to different dialects (Flege and Bohn 2021). Individual differences in L1 categories influence how phonetically dissimilar from L1 categories speakers perceive L2 sounds. This was demonstrated by Chládková and Podlipský (2011), who found that Moravian Czech listeners assimilated the phonetically short Dutch vowels /i, ɪ/ to their short L1 category /i/ based on duration, while Bohemian Czech listeners used their two L1 categories (/i:, ɪ/) to differentiate the vowels, relying on spectral cues rather than duration. The study concluded that Dutch learners would follow different paths in acquiring these two vowels, depending on their L1 dialect. The findings underscore the necessity to refer to individual learners' L1 categories, also considering their dialect background.

2.3 Cross-language assimilation patterns

Focusing on learners who already have a phonetic system for their L1 at the onset of L2 acquisition, PAM-L2 considers the type of perceptual assimilation of two contrasting L2 phones to one or more L1 categories determinative of learners' success in discriminating L2 contrasts at least during the early stages of L2 learning (Best and Tyler 2007). Here, the L1-L2 vowel mapping pattern is used to predict both the learners' perception and production, though an individuals' categories used in perception and production do not have to agree (Flege and Bohn 2021).

According to PAM-L2, discriminating between two distinct L2 sounds is not difficult when each phone is mapped onto a different L1 category (two-category assimilation), which eliminates the need to learn the L2 contrast, as the learner's established L1 categories are sufficient for discrimination of the L2 sounds. When the L1 and L2 sounds are perceived as similar enough, a shared phonetic and phonological L1-L2 category is used for them (Tyler 2019). The learner will not produce a difference between the phones in L1 and L2 in the case of a two-category assimilation if he uses a common L1-L2 category (Best and Tyler 2007).

According to PAM-L2 (Best and Tyler 2007), if two distinct L2 sounds are mapped onto the same L1 category (one-category assimilation), discrimination will be difficult. At least initially, the learner will not perceive a contrast between the two L2 sounds, so perceptual learning will be necessary for the contrast to be discriminated. The two sounds may not be produced distinctively, instead, they will be produced as the L1 sound they map onto. Nevertheless, the learner may differentiate between the L2 sounds as one being a more appropriate example of the L1 category than the other (category-goodness assimilation). A new category is most likely to be established for the sound that is perceived as the less accurate example of the L1 category, allowing for the two L2 sounds to be produced contrastively in time.

Regarding the vowels studied here, the learners are expected to assimilate the L2 vowels /i-ɪ/ to the L1 /iː-ɪ/ in a two-category assimilation. The preservation of the quality contrast depends on the differentiation of the corresponding L1 vowels in terms of quality. As the participants in this study come from Moravia, they are expected to differentiate the L1 vowels primarily in quantity, with quality being only a secondary cue (Podlipský, Skarnitzl and Volín 2009). The learners are likely to produce the high front Czech vowels as close to one another in quality, though they may be somewhat differentiated, with /ɪ/ being the lower and more retracted of the two. As the English vowels /i-ɪ/ are expected to be contrasted in primarily duration under the influence of the learners' L1, the learners are not expected to clearly separate them in quality, if their L1 vowels are not clearly separated in quality. With the quantity cue reused, /i/ will be realised as the longer of the two vowels, allowing the learners to use two composite L1-L2 categories: one for /i, iː/ and another for /ɪ, ɪ/. The learners are not expected to improve in producing the two categories contrastively in quality, as implementing a qualitative difference and discarding the use of quantity as contrastive is unlikely due to the prevailingly L1-accented input and little L2 experience.

In this study, /æ/ may be involved in two different assimilations based on the children's source of input. The teacher realised /æ/ as relatively open, and so the learners are expected to perceive /æ/ as close to /a/ in quality, whereas non-native input outside lessons is assumed to have provided /æ/ tokens that are realised as the Czech /ɛ/. In the case that /æ/ is mapped onto /a/, the /æ-ʌ/ contrast will form a category goodness assimilation to the Czech /a/, with /æ/ being the less acceptable instance of the L1 vowel. If /æ/ is mapped onto the Czech /ɛ/, the /ɛ, æ/ will not be produced as different in quality. In the input that does not preserve the /æ-ɛ/ contrast, the two vowels are not expected to be differentiated in quality, hence this predicted lack of differentiation does not arise from an assimilation of two L2 sounds to one L1 sound, but is caused by the nature of the input, which merges the two vowels. Additionally, as learners may reuse vowel quantity to differentiate L2 vowels, they may also contrast the vowels /ɛ, æ/ or /æ, ʌ/ primarily in duration, failing to notice or implement the quality difference.

According to Best and Tyler (2007), if a learner does not perceive a difference between two L2 sounds in category-goodness assimilation, both L2 sounds will function as phonetic variants of the same L1-L2 phonological category. With exposure and practice, if the learner notices that the two L2 sounds signal a meaning difference between minimal pairs, a new phonological category may be established for the sound that is perceived as the less acceptable instance of the L1 sound. The sound considered the more acceptable instance of it will be perceived as phonetically and phonologically equivalent to that L1 category.

Table 1. L1-L2 vowel category assimilation type, mapping pattern, and category used in the L2 for each English vowel studied.

Assimilation type	**L2 vowel**	**L1 vowel**	**Category used in L2**	**Expected L2 contrast**
Two-category assimilation	i	iː	Common L1-L2 category	Quantity contrast
	ɪ	ɪ~i	Common L1-L2 category	
Assimilation of (1) /ɛ~æ/ to /ɛ/; or (2) category goodness assimilation of /æ-ʌ/ to /a/	ɛ	ɛ	Common L1-L2 category	(1) Quantity contrast: /ɛ/ as the shorter vowel, quality contrast between /æ/ and /ʌ/; (2) quantity contrast: /ʌ/ as the shorter vowel, quality contrast between /æ/ and /ɛ/; (3) quality contrast between all 3 vowels
	æ	ɛː aː	(1) /ɛ/-like /æ/, or (2) /æ/ as a phonetic variant of /a/; or (3) /æ/ as a new phonological category	
	ʌ	a	Common L1-L2 category	

Šimáčková and Podlipský (2018) and Šturm and Skarnitzl (2011) show that the /æ-ɛ/ contrast poses difficulties even for experienced Czech learners. Šimáčková and Podlipský (2018) investigated production of L2 vowel pairs /i-ɪ/ and /ɛ-æ/ by Czech university students of English. The students maintained a contrast between the two high front vowels in both height and retraction. However, the /æ-ɛ/ pair was primarily contrasted in duration, with /æ/ realized as the longer vowel. The students did separate the pair in height somewhat, but lexical misrepresentation was common. The results indicated that the learners perceived sounds belonging to the /æ/ category as less acceptable instances of /ɛ/.

Šturm and Skarnitzl (2011) examined Czech learners' (secondary school and university students) perception of the English /æ/. Learners with more L2 experience were more accurate and consistent in categorising /æ/ compared to those with less experience. The experienced group's judgments were influenced by vowel height and F1-F2 ratio, whereas the less experienced learners were inconsistent in their judgments of /æ/ tokens and did not attend to vowel height, overlapping /æ/ with /ɛ/. However, no significant correlation was found between the duration of /æ/ and acceptability in either group of hearers.

2.4 Non-native input and exemplar theory

The possibility of /æ/ being mapped onto different L1 categories depending on the production of /æ/ by a given source of input, and the impact of the different resulting mapping pattern on the learners' production is explored in this section. High-quality native speaker input is crucial for an optimal EFL learning experience (Tyler 2019), as learning to produce phonetic categories absent in the L1 is greatly influenced by the quality of input (Escudero 2001). Foreign accented input does not necessarily inhibit the development of a learner's ability to preserve L2 contrasts. As long as all L2 phonological distinctions are maintained and perceived correctly, the input can facilitate the acquisition of L2 phonological contrasts (Tyler 2019), even though learners are unlikely to produce them in a native-like manner (Flege and Eefting 1987). However, further complicating the issue, learners may be exposed to multiple sources of input, which differ in how they specify L2 vowels. High input variability has been reported to lead to variable vowel production in the L1 (Levy and Hanulíková 2019), as it can cause difficulty with forming stable sound representations for L2 vowels (Bosch and Ramon-Casas 2011).

In the present study, the learners received interactional input from two major sources. Input provided by the teacher during lessons preserved the relevant L2 contrasts. The other major source of input was provided by the learners' family members, who self-reported that they exhibit strong Czech foreign accent. A common feature of Czech-accented English is the lack of preservation of the /æ-ɛ/ contrast (Skarnitzl and Rumlová 2019). It is assumed here that input outside of lessons did not generally distinguish between the two vowels qualitatively. The dual input is expected to affect the production of L2 vowels, and the nature of this effect may be accounted for with the use of Pierrehumbert's (2001) exemplar-based model. Concerning the other vowels examined, the vowels /i, ɪ/ are predicted to map onto the two Czech high vowels based on quantity irrespective of the input source (unlike /æ, ɛ, ʌ/), thus, this vowel pair is not expected to be more or less differentiated based on input source.

Exemplar-based models like that of Pierrehumbert (2001) posit that for every perceived variant of a linguistic item, an exemplar (unique representation) is stored within an exemplar cloud in the listener's mind. The exemplar cloud encompasses all the different exemplars of the given item. Similar exemplars are located close to each other, dissimilar ones far apart. Storing every encountered item as a separate exemplar means that categories whose instances are encountered more frequently will have a larger number of exemplars associated with them, while categories that are encountered less frequently will have fewer associated exemplars. According to Pierrehumbert's (2001) model, phonetic details may not only be associated with languages or dialects, but also with individual words. The lexicon and grammar are considered two generalizations over the same memories, allowing for variations in pronunciation to be encoded in lexical entries, and for exemplar clouds to occur concurrently at the segmental and lexical levels within a learner's mind.

Pierrehumbert (2001) extends his model to also account for production. Exemplar-based categorisation uses the encoded phonetic features in perceived words to localise similar exemplars in the parameter space and save the new exemplar among them. For production, he probability of a specific exemplar being chosen is directly related to its activation. Exemplars that have been recently and frequently activated attain full activation and consequently have a stronger influence on production.

Due to the different assimilations /æ/ participates in based on source of input, instances words containing /æ/ are expected to be correspondingly divided into two groups depending on whether a given word has more exemplars provided by input from the teacher, or input outside lessons. If the learners associate encountered L2 sounds with the

L1 category they map onto, /æ/ word tokens heard outside lessons will be stored as word exemplars containing a vowel that is the same as the Czech /ɛ/, while /æ/ word tokens provided by the teacher will be stored as word exemplars containing an /a/-like vowel.

This categorisation will then be reflected in dual production of words: as exemplars of /æ/ words encountered (1) outside lessons are likely to (a) be heard in old words revised outside class, and (b) contain an /ɛ/(-like) vowel (rather than an /a/-like vowel), old /æ/ words should therefore activate word exemplars with an /ɛ/(-like) vowel, causing production to be biased towards producing /ɛ/ in old words that actually contain /æ/. Tokens of /æ/ heard (2) during lessons, on the other hand, will be realised (a) in new words recently taught during lessons, and (b) as more /a/-like, leading to a bias towards /a/-like production in new words.

The division of words into old and new, however, does not reflect a clear-cut distinction. The learners also heard some old words from the teacher, though probably not as often as outside class. The teacher's productions of old words are likely to cause confusion, which could, however, help with noticing the similarities between old and new words, and thus aid in separating /æ, ɛ/ in the old words. However, words encountered more frequently have more numerous exemplars with higher activation, leading to denser and more activated exemplar clouds. Therefore, in situations where there is ambiguity, a bias towards the more frequent label is predicted. So, in the case of old word tokens provided by the teacher, the learners' production is expected to still be biased towards producing /ɛ/, as /ɛ/(-like) vowel exemplars are expected to be more frequent and more activated.

2.5 The effect of L2 vocabulary size on L2 learning

Though improvement in learners' perception of difficult L2 contrasts through perceptual learning is likely in immersion settings (Wang et al. 2023), the probability that the learners will acquire a new L2 category is generally lower in classroom settings due to the limited opportunities for perceptual learning of L2 phonological contrast before the establishment of an extensive L2 vocabulary (Tyler 2019; Bundgaard-Nielsen, Best and Tyler 2011). Rapid acquisition of a large L2 vocabulary is hypothesised to interfere with perceptual learning because of perceptual adjustment occurring during the early stages of language acquisition, before the establishment of the lexicon and higher-order linguistic structures (Best and Tyler 2007). Hence, L2 learners should be provided with learning experiences focused on L2 phonological contrasts at the earliest stages of L2 learning, ideally before a "large" L2 vocabulary is established (Tyler 2019).

The learners studied here were acquiring L2 vocabulary at a rapid pace, and when the experiment was conducted, all have surpassed the 50-word mark, which Bundgaard-Nielsen, Best and Tyler (2011) suggests to be the approximate point where the vocabulary becomes "large." Even the least experienced learner (speaker 1), with 5 months of learning, is expected to have already acquired around 300 words by start of the experiment (Helen Doron English 2022) The more experienced children who had been attending lessons for more than 2 years, would then be expected to have learned about 1600 words prior to the experiment.

A large increasing vocabulary may aid perceptual learning for the contrasts that are relatively to discriminate (like single-category assimilations), while inhibiting perceptual learning for those that are difficult, like single-category assimilations (Bundgaard-Nielsen et al. 2011). Increasing vocabulary that the children are building reinforces the L2 contrasts they have acquired, as well as possibly reinforcing any category goodness assimilation patterns they may have established. The reinforcement of an /ɛ/-like L1-L2 category for vowels in words containing both /æ/ and /ɛ/ based on input that provides old word tokens is likely to lead to misrepresentation of words containing /æ/ as containing /ɛ/. This kind of lexical misrepresentation has even been reported for highly proficient Czech-English bilinguals (Šimáčková and Podlipský 2018).

The ideal input targeted at perceptual learning of the /æ-ɛ/ contrast should establish a vocabulary slowly, utilising words that demonstrate the contrast. In the present study, teaching words with these two vowels at this stage of the learners' L2 learning may still help them notice the differences between the vowels, as the teachers' realisation of /æ/ is noticeably /a/-like, even if the learners notice them only in the new words. However, an increasing vocabulary focused on the /i-ɪ/ contrast may reinforce the quantity-based distinction between the vowels, not leading to improvement in contrasting them in quality.

2.6 The impact of children's language development on L2 vowel acquisition

Trends in child language development influence vowel production in both L1 and L2. Children's vowel productions are known to be more variable than those of adults, with formant frequency variability reaching adult levels only around age 14 (Lee, Potamianos and Narayanan 1999), making substantial category overlap likely in productions of young children in both their L1 and L2. Walley and Flege (1999) found that five-year-olds' phonemic categories to be wider than those of nine-year-olds, demonstrating that, as children mature, their ability to discriminate L1 sounds improves. Thus, phonetic representations in children's L1 are frequently not entirely adult-like. Best and Tyler (2007) also note that children around the third and fourth year of life are still refining their perception of L1 sounds, and perception in children only reaches adult-like levels for L1 speech by 8 years of age.

Simon, Sjerps and Fikkert (2014) found 9–12-year-old EFL learners to have underspecified L2 vowel representations, especially for those vowels which participated in contrasts absent in their L1. The children accepted more incorrect pronunciations of L2 words (based on vowel substitutions) than adults, which was taken to indicate that for L2, the children were using variable vowel categories that encompassed neighbouring vowels. Baker et al. (2008) also report that children tend to use broad categories in perception, resulting in a large overlap between neighbouring vowel sounds. Young children's vowel categories in both production and perception are expected to be variable, and especially their vowel categorisation in L2, considering that they are inexperienced learners, cannot be expected to be very accurate. As child learners' shared L1-L2 phonological system continually changes with experience, their production and perception of L1 and L2 sounds become more adult-like (Walley and Flege 1999, Lee, Potamianos and Narayanan 1999), and can also affect one another in different ways. The influence that L2 learning may exert on how L1 sounds are produced is dealt with below.

2.7 Phonetic drift in child learners

L1-L2 interactions are bi-directional, with the possibility of L2 categories affecting L1 categories (Flege 1987, Chang 2012, Kartushina 2016). This "phonological restructuring of L1 as a consequence of L2 experience" is referred to as phonetic drift (Chang 2012, 249). Chang (2012) proposed that phonological similarity perceived between L1 and L2 sounds leads to L1 drift in inexperienced learners, which is in opposition to Flege (1987), who found stronger L1 drift in experienced rather than in inexperienced learners. L2 categories assimilate (Yang and Fox 2017) or dissimilate (Yusa 2010; Flege, Schirru, and MacKay 2003; Guion 2003) from the closest L1 categories.

Chang (2012) reported that after just six weeks of intensive Korean (L2) immersion, native English speakers showed assimilatory drift of their English (L1) sounds. Observing that inexperienced Korean learners exhibited significant shifts in their production of English stops and vowels, while experienced Korean learners showed less dramatic drift, it was proposed that this pattern may be a result of the novelty effect that occurs when learners encounter a new L2 sound system. Similarly, Kartushina et al. (2016) reported that after receiving short-term visual articulatory feedback training, native French speakers showed a drift of their native vowels towards non-native (Russian) vowels.

Investigating L1 drift exhibited by child learners, Yusa et al. (2010) focused on VOT production of Japanese children learning EFL in kindergarten (age at onset of learning = 4 years old) with a native speaker, comparing their L1 and L2 stop productions after 20 and 32 moths of learning. After 20 months, the learners with more L2 exposure (hours of daily interaction in English) exhibited dissimilatory VOT patterns. Learners who interacted with the native speaker only occasionally (less than 30 minutes daily) employed L1 VOT values. After 32 months of exposure, even this group showed dissimilatory drift. It was hypothesised that in the initial phase of L2 learning, where learners are most dependent on their L1, the L1 may be more subject to modifications than L2, suggesting an L2 effect on the L1 of learners who are not proficient. This study further implies that because the learners had little exposure to the L2, they remained in the initial phrase of L2 learning for years after they began learning.

Both Chang (2012) and Kartushina et al. (2016) suggest that native vowel drift in adult learners occurs during the initial stage of L2 learning. However, some studies of learners differing in L2 experience report results contradictory to those of Chang (2012). Yang and Fox (2017) found no evidence to support the idea that L1 vowel production was affected by L2 learning in child learners with low L2 experience (average of 3.9 months' residence in the US). On the other hand, the more experienced children (almost three years of daily exposure to English in an immersion setting) showed noticeable changes in L1 vowel quality, indicating assimilatory movement. In this, the results are contradictory to those of Chang (2012), but in line with Flege (1987), who also reported stronger L2 effects on L1 in learners with more L2 experience.

The context of L2 learning differs dramatically between the studies mentioned, making comparison and formulation of predictions difficult. Based on Chang (2012) and Kartushina (2016), who show that a drift effect occurs during early stages of L2 learning in adults, and Yusa (2010), who shows a drift effect for low-proficiency EFL child learners with 32 months' of exposure to English, a drift effect can be expected also for the children studied here. However, Yang and Fox (2017) found no drift in inexperienced children learning immersively. As the study by Yusa (2010) focused on participants most similar to those in the present study (EFL learning in an L1-dominant environment, length of exposure around 2.5 years), based on its findings, it can be expected that the children in the present study will also show L1 drift.

2.8 Summary of the predictions

So far, there have only been a few studies focusing on L2 vowel production in child EFL (Yusa 2010), and no research has yet been conducted on the development of English L2 vowel contrasts among Czech children. The main objectives of this study are twofold: first, to investigate the extent to which preschool EFL learners differentiate contrasting L2 vowels in speech production; and second, to explore if the process of learning an L2 affects the production of L1 vowels. More specifically, the study focuses on three questions. (1) Is the extent to which the learners differentiate L2 vowels affected by whether their source of input preserves L2 contrasts, and if so, how? (2) Do the learners show improvement in producing L2 vowels contrastively with time? (3) Is L2 learning accompanied by changes in production of L1 vowels? Additionally, the research examines L2 vowel perception in words and non-words.

To summarise the predictions formulated in previous sections, the L1-L2 vowel mapping pattern based on Best and Tyler (2007) and the learners' source of input are seen as potentially affecting the learners' production of L2 vowels. Table 2 recapitulates the predictions for production. Based on input heard outside lessons, the vowels /æ, ɛ/ are expected to map onto the Czech /ɛ/, as this input is assumed to realise both the L2 vowels as the same as the L1 /ɛ/. Based on lesson input, /æ, ʌ/ are expected to map onto the Czech /a/, with /æ/ as the less acceptable instance of /a/. Based on Pierrehumbert (2001), as a consequence of their dual input, the learners are expected to produce /ɛ, æ/ as more differentiated in quality in words that they have acquired during lessons (new words), and had not had the chance to practice outside lessons. On the other hand, the words they practiced outside lessons (old words) are expected to be represented by more Czech-accented exemplars, leading the learners to not produce /ɛ, æ/ as separated in quality. The /æ-ʌ/ contrast is expected to be separated in quality in old words, but may be separated less in new words, due to /æ/ being produced as lower under the influence of lesson input. As for perception, it was tested only for old words, so /æ, ɛ/ were expected to be discriminated poorly, like in Šturm and Skarnitzl (2011).

The English /i, ɪ/ are expected to map onto the Moravian Czech /iː, ɪ/ in a two-category assimilation. The learners are, however, expected to reuse the L1 quantity cue as the primary cue used in discriminating the L2 vowels, like the learners in Šimáčková and Podlipský (2018). Consequently, the English vowels are expected to be produced as sufficiently different in quality if the corresponding Czech vowels are also produced as sufficiently different. Input source is not expected to affect the production of these vowels, as the two-category assimilation based on duration is likely to help the learners discriminate the vowels relatively well, considering that their non-native input may also preserve a quantity contrast between the two, and that they have little exposure to the L2. As for perception, if the quality difference is not attended to, perception will not be accurate, as the consonantal context was not controlled in the task, and so the quantity cue may not be reliable.

Regarding the effect of time, as data from six English production sessions were split into only two times, it is unlikely that time will significantly affect L2 vowel production. Improvement in the separation of /ɛ, æ/ in old words is possible, but not very likely, as based on Pierrehumbert (2001), input outside of lessons is expected to provide more exemplars of the old words than lesson input throughout the whole experiment. Likewise, an improvement in separating /i, ɪ/ in quality, if they are not sufficiently separated during the first six weeks, is not expected, due to learners contrasting the vowels in length in a two-category assimilation.

Based on the limited research conducted on preschool EFL learners, phonetic drift of L1 categories can be expected, this prediction is formed in accordance with Yusa (2010), who reported drift in children who had been attending EFL lessons for more than 2.5 years, similarly to most children in this study. Chang (2012) and Kartushina (2016) also support the expectation that low-proficiency learners will show L1 drift.

Table 2. L1-L2 vowel mapping pattern based on the source of input.

Source of input	Word status	L2 vowel	L1 vowel	Expected production
Prevailingly input outside lessons	Old words	/i/	/iː/	Differentiated primarily in duration
		/ɪ/	/ɪ~i/	
		/ɛ/	/ɛ/	/ɛ-æ/ differentiated in duration,
		/æ/		
		/ʌ/	/a/	/æ-ʌ/ differentiated in quality
Lesson input	New words	/i/	/iː/	Differentiated primarily in duration
		/ɪ/	/ɪ~i/	
		/ɛ/	/ɛ/	/æ-ɛ/ differentiated in quality,
		/æ/	/a/	/æ-ʌ/ differentiated in quality
		/ʌ/		

3 Methods

3.1 Participants

A total of nine Czech learners of English aged 3;9 to 5;9 participated in the study. Additionally, 10 age-matched Czech monolinguals were recorded to serve as monolingual reference points to the learners. The English learners were attending 45-minute once-weekly EFL classes at the time of the experiment. The group of learners included seven children (six girls and one boy) of ages 5;4 to 5;9 (Mean age = 5;6, SD = 2 months), one girl aged 4;8, and one girl aged 3;9. On average, the five-year-olds had been attending the EFL lessons for longer (Mean length of exposure to English = 32 months, SD = 4.9 months) than the younger children. The age at which the learners started learning English ranged from 1;9 to 4;9 (Mean = 3;0, SD = 7.6 months).

The participants were attending beginner courses, targeting learners at the proficiency level A1 according to CEFR (Council of Europe 2001). Before being enrolled in the EFL classes, the learners had little to no exposure to English, as reported by parents. Czech was the dominant everyday language used in all participants' families. In a questionnaire (based on Blumenfeld and Kaushanskaya 2007, see the Appendix), the families reported Czech as the only (or largely dominant) language used at home and in the children's kindergarten. All were raised in a near-monolingual environment, except speaker 8, who was a Czech-Slovak simultaneous bilingual, with Czech as her dominant language; and speaker 2 was marginally exposed to Belarusian. The dialectal background of the learners was similar, with speaker 2 reported to have more experience with Bohemian Czech than the other learners, but all (including speaker 8 and 2) were exposed to Moravian Czech predominantly.

Outside their EFL lessons, most children actively practised speaking English on a weekly basis at home and had a positive relationship to English and their EFL lessons. All class-related audio materials, online games and short videos featured exclusively speakers of SSBE; while internet videos, TV programmes and mobile games that the children were exposed to at home often included speakers of GA. The teacher was a highly proficient native Czech speaker of L2 English with an accent approximating SSBE. The only source of reciprocal interaction outside lessons were parents, who rated their accents in English as influenced by Czech to a large degree.

All participants were typically developing children showing age-appropriate oral-motor skills. Their articulation was assessed (Soubor orientačních testů 2007), as well as their passive and active L1 vocabulary (Vidnerová 2014). No participant was reported to have any hearing, speech, or language problems.

Table 3. Characteristics the learners: subject number, sex, age, length of exposure to English in months, age at onset of EFL learning.

Subj.	Sex	Age	Length of exposure	Onset of EFL learning
1	F	3;9	5	3;4
2	F	4;8	6	4;2
4	F	5;4	30	2;10
5	F	5;4	30	2;10
6	F	5;4	43	1;9

7	F	5;6	30	3;0
8	F	5;7	31	3;0
9	F	5;9	30	3;3
10	M	5;8	30	3;2

3.2 Reference data

The group of Czech monolinguals consisted of 10 children (seven girls and three boys) corresponding in age to the learners (see the Appendix for participant details). All learned Czech as their L1 had little to no knowledge of English or any other language than Czech. Their dialectal background was similar to the learners'.

Additionally, YouTube videos of eight age-matched British children who spoke SSBE were used to for comparison with the learners' English vowels. The YouTube videos were conversational interviews, in which the child was questioned by their mother (see the Appendix for participant details).

3.3 Stimuli

A picture-naming task was employed to elicit target vowels. The targets for analysis were the five SSBE vowels /i, ɪ, ɛ, æ, ʌ/ in 39 predominantly monosyllabic words of /CVC/ structure, or disyllabic words with the stressed target vowel in the first syllable. The adjacent consonants were most often obstruents, though some words featured a prevocalic approximant (see Appendix for a list of all words). The English words were divided into those known to the learners at least for a month at the time of the first recording (old words) and those recently learned during lessons (new words). 16 Czech words containing the vowels /iː, ɪ, ɛ, a/ recorded in two sessions separated by 10 weeks. All words were selected based on imageability.

3.4 Procedures

3.4.1 Recording procedure: Production sessions

Participants produced the words spontaneously, if they could not remember a word, a delayed imitation procedure was used (the word was repeated after the teacher). The data collection for the learners was completed in the span of three months in a classroom of a language school where the English learners' lessons took place, with the recording session always immediately following the English lesson. Each new word was recorded in two sessions, nine of the old words (*back, cat, hat, head, bed, leg, bus, fish, sheep*) were recorded in every session, and the remaining old words were only recorded in one session.

For the learners, there were a total of eight production recording sessions: six English sessions and two Czech sessions. English production sessions were separated by a three-week period (see Appendix for detailed plan of recording sessions). The teacher interacted with the participants in Czech during all sessions to ensure that they understood instructions. During the recording session, the participant was seated in a quiet room inside a tent facing a laptop. A cotton canvas tent was used to eliminate echo in the room. The teacher was sitting outside the tent. Verbal praise was used during the session to motivate the children to continue with the session. A set of pictures representing target words was presented on the laptop screen in random order. The number of pictures shown per session varied from 16 to 21. After the pictures had been shown once, the child either chose to see the set again or finish the session. In most cases, the set was shown twice, in a different random order each time.

Seven of the monolingual Czech children were recorded at a kindergarten near Olomouc, three were recorded at their homes. The procedure with the Czech monolinguals was done over two consecutive days, with the Czech vocabulary and pronunciation assessment being done on the first day just before session one, session two being conducted on the following day. They were recorded in a quiet room, sitting inside the same tent. During the recording session a H4n zoom recorder was held by the experimented approximately 30 cm from the participant's mouth. Speech samples were recorded with 16-bit quantization rate and 44.1 kHz sampling rate.

3.4.2 Recording procedure: Perception sessions

Each English learner attended two English perception sessions, which targeted the learners' ability to recognise mispronunciations of English words. In each session, 12 mono or disyllabic words were produced by the teacher, each either with the correct stressed vowel (in the first syllable), or with a replaced L2 vowel.

During the session, the teacher was seated in the tent holding a dog puppet next to a laptop. The teacher produced the words her natural voice, pretending the puppet is pronouncing them. The corresponding picture for each word was displayed on the laptop screen, the pictures were shown in random order, with the fillers always occupying the seventh and the final slot (see Table 4 for vowel substitutions). After the word was spoken, the child reacted with "no" to productions perceived as incorrect (to indicate that the picture did not match the uttered form) and "yes" to words perceived as uttered correctly. The sessions were separated by three weeks for most learners, and by one week for speaker 3. The setting was the same as for the production sessions. The teacher recorded her production holding the H4n zoom recorder approximately 20 cm from her mouth. The recorder settings were kept identical to those used during the production sessions.

Table 4. Perception test words. Fillers are italicised. Asterisk indicates control vowel change.

Test words	Possible produced vowels	
	Correct vowel	Replaced vowel
cat, back, rat, hat, rabbit, carrot	æ	ɛ, ʌ
yellow, leg, bed, hen, shell, cherry	ɛ	æ, ɛ, ɪ*
fish, biscuit, chick, kitten	ɪ	i
feet, sheep, tree, knee	i	ɪ
goose, shoes	u	-
scarf, grass	ɑ	-

3.4.3 Teaching procedure

The English learners were taught all words during their EFL lessons using the Helen Doron teaching approach (Doron 1993), adhering to its lesson plans. The old words, taught at least a month before the experiment began were revised in at least one class preceding the production session. New words, taught during the experiment, were included in the recording session immediately following the lesson in which they were taught, revised in the two following lessons, and recorded again three weeks after they had been first taught (see Appendix for detailed plan of word revision and teaching).

The basic features of the HD approach are: L2-only interaction during lessons (Sheffler 2015) use of repetitive speech structures (Doron 1993), meaningful input provided in task-based learning, verbal encouragement, and exposure to L2 audio outside lessons (Doron 2011). The teaching approach stresses frequent (twice daily) and plentiful at-home listening to "texts, songs and rhymes" (Doron 1993, 4) that tie into the material taught during lessons (Sheffler 2015). Children learn the meaning of words and phrases from the audio via meaningful input provided during their in-class task-based learning (Sheffler 2015). Repetitive speech is used and encouraged by the teacher to promote memorization of "lexical items and grammatical structures" (Doron 1993, 3). It is coupled with "positive reinforcement" (Doron 1993, 3) to motivate the children (Doron 2011). "Demonstrative praise" is used as reward for the child's participation in conversation (Doron 1993, 3).

HD carries many of the features of Communicative Language Teaching (see Dörnyei 2009 for a discussion of CLT), and incorporates Total Physical Response (Asher 1969). Its aim is to provide meaningful input from simulated daily tasks (Kovács in Doron 2011), which are reproduced using "visual, auditory" and "kinaesthetic" tasks (Helen Doron English 2022). The teacher acts more like a playmate to the child, than an authoritative figure (Doron 2011).

The specific lesson tasks are contests, role play, dancing (Doron 2011), "colouring" (Doron 1993, 4), "drawing" and "conversation" (Klimova 2013, 504), various game activities (Doron 2011) that incorporate learning aids (Doron 2011) like toys (Doron 2011). The tasks cover topics like "greetings, family members, numbers, colours, animals, house, body parts, food, fruit and vegetables, clothes, weather" (Klimova 2013, 504), using vocabulary appropriate for beginner learners (Doron 2011). Each learner group consist of four to eight children (Tefelnerová 2013, 48).

3.4.4 In-class perception trials

A single in-class perception trial was conducted in each lesson to prime the children for the perception tests. In-class perception trials were short activities (less than a minute) in which the teacher presented the children with four flashcards familiar to them and pronounced the name of the depicted object with the correct stressed vowel or a replaced stressed vowel. The words included only non-target vowels (back vowels and diphthongs). A puppet was used while uttering the words. Children collectively reacted "yes" or "no," as in the perception tests.

3.5 Measurement and Analysis

3.5.1 Formant and F0 measurements

Following Machač and Skarnitzl (2009), vowel onset and offset points were marked by hand and determined primarily on the basis of visible F1 and F2 tracks present in the spectrogram. The segments used for formant and pitch measurements were 25-50 ms segments that showed stable F1 and F2 tracks (Derdemezis et al. 2016) located within the larger whole vowel segments. Other criteria used to determine the location of the segment were F3 stability, intensity, and proximity to midpoint. Where LPC tracks were judged to be inaccurate, the measurement was done manually using visual information present in the spectrogram.

Measurements of f0, F1 and F2 were done in the same way for the English learners, Czech monolinguals, and English monolinguals from YouTube videos. A modified script (de Haro, 2022) was used to extract the mean F1, F2 and f0 values over the 25-50 ms segments in Praat (Boersma and Weenink 2022). The maximum number of formants was set to 5 and maximum formant frequency to 7000 Hz (Viegas, Viegas and Baeck 2015). The time step was adjusted to 0.00625 (Derdemezis et al. 2016); minimum pitch was set to 100 Hz and maximum pitch to 800 Hz (Viegas, Viegas and Baeck 2015).

3.5.2 Statistical analysis

For the analysis, the six English sessions were split into two times: sessions 1 through 3 as time 1, and sessions 4 through 6 as time 2, to reduce the complexity of the analysis and account for learners' absences. The line between time 1 and time 2 was thus drawn arbitrarily, with time 1 encompassing productions from the first six weeks of the experiment, and time 2 encompassing productions from the last five weeks, with a three-week interval separating the latest time 1 productions from the earliest time 2 productions. Because the six sessions were separated into two times, the effect of time is not expected to have a drastic influence on vowel production when it comes to the English vowels. However, for the Czech vowels, time 1 and time 2 correspond to the two Czech sessions, which were separated by 10 weeks.

The raw formant frequency values in Hz were ERB-normalised using the R package phonR (McCloy 2016), which employs the method from Glasberg and Moore (1990). The data were analysed in R (R Core Team 2023), using the packages lme4 (Bates et al. 2015), emmeans (Length R 2023), phonR (McCloy 2016), and MuMIn (Bartoń 2022) (see the Appendix for the analysis script).

4 Results

Two pairs of linear mixed-effects models were fitted in R using the lme4 library. One pair of models estimated height and retraction of English vowels spoken by the learners in old and new words at time 1 and 2. The other pair was used to compare height and retraction of Czech vowels produced by seven of the learners at time 1 and 2. Vowel height was predicted as f0 subtracted from F1 (in ERB), and the other to predict vowel retraction as f0 subtracted from F2 (in ERB). All models were estimated by REML using the BOBYQA optimizer. P-values were computed using Satterthwaite's method, and the alpha level was set to 0.05. It is assumed that the response data was generated by a Gaussian process.

4.1 English vowels produced by the learners

For the English target vowels spoken by the nine learners, two models were used to estimate height and retraction as a function of the fixed effects *vowel*, *time*, and *status*, including a three-way interaction between them. Both time and status were specified as treatment coded two-level categorical predictors, with time 2 and status "new" as reference levels. Vowel was specified as a factor predictor with five levels: the target vowels /i, ɪ, ɛ, æ, ʌ/ with /æ/ as the intercept. The random effects estimated by the model were *speaker*, *word*, and *time*. By-word varying intercepts were estimated, as well as by-participant varying slopes for time. Across the five vowels, 971 productions (word tokens) were supplied to the model.

4.1.1 Vowel height

The model for learners' L2 vowel height was estimated using the formula *F1e-F0e ~ vowel * time * status + (1 + time | speaker) + (1 | word)*. The total explanatory power of the model was substantial (conditional $R^2 = 0.71$), with the fixed effects alone accounting for 64% of the variance described by the model (marginal $R^2 = 0.64$).

The vowel /æ/ in the intercept (in time 2, status "new") was estimated to be 8.29 (SE = 0.378, t = 21.912, p <0.001), meaning that the height of /æ/ productions in time 2 in new words was estimated to be 8.29 ERB. The values estimated for other are reported relative to /æ/ (at time 2 in new words). Importantly, the coefficient table shows a significant p-value for /ɛ/ ($\beta = -2.031$, SE = 0.395, t = -5.143, p < 0.001), indicating that productions of /ɛ/ differed in height from /æ/ in new words at time 2. This is also the case for /ʌ/ ($\beta = -0.902$, SE = 0.446, t = -2.025, p = 0.046), meaning that /æ/ and /ʌ/ were separated in height in new words at time 2 (Table 5 shows the model's coefficient table). As the vowels /æ, ɛ, ʌ/ were not produced as significantly different at time 1, they can be assumed to have been similarly separated at time 1.

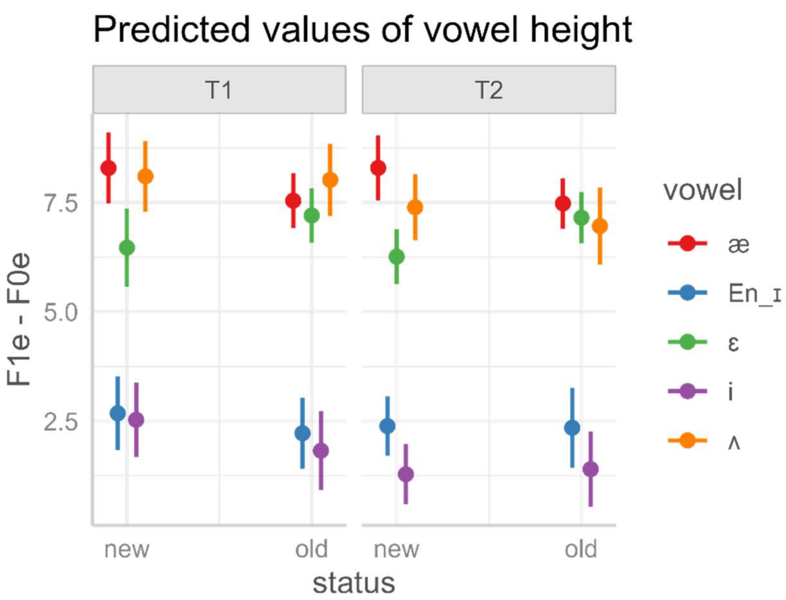
The effect of time was significant only for /i/. Productions of /i/ differed in height between time 1 and time 2 (in new words) enough for the shift to be significant ($\beta = 1.247$, SE = 0.597, t = 2.089, p = 0.038). The slope estimate was positive, meaning that /i/ was produced as more closed in time 2 in comparison with time 1 (see Figure 2).

The effect of word status "old" on /æ/ productions is negative and significant ($\beta = -0.813$, SE = 0.374, t = -2.175, p = 0.034), indicating that /æ/ was produced as more open in new words when compared to /æ/ in old words. The vowel /ɛ/ also differed significantly in height depending on word status ($\beta = 1.704$, SE = 0.489, t = 3.485, p = 0.001), the positive slope indicates that /ɛ/ was produced as more closed in new words when compared to old words (see Figure 1), indicating a better separation of /ɛ/ and /æ/ in new words.

Table 5. The coefficient table for predicted vowel height of the L2 English vowels /i, ɪ, ɛ, æ, ʌ/ produced by the learners. The vowel in the intercept is /æ/. The reference level for time is time 2, for status, the reference level is new words.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	**8.293**	**0.378**	**44.031**	**21.912**	**< 0.001**
ɪ	**-5.906**	**0.416**	**64.137**	**-14.183**	**< 0.001**
ɛ	**-2.031**	**0.395**	**65.667**	**-5.143**	**< 0.001**
i	**-7.007**	**0.42**	**67.668**	**-16.688**	**< 0.001**
ʌ	**-0.902**	**0.446**	**74.283**	**-2.025**	**0.046**
timeT1	-0.003	0.469	82.522	-0.006	0.995
status2	**-0.813**	**0.374**	**59.162**	**-2.175**	**0.034**
ɪ:timeT1	0.295	0.596	129.592	0.495	0.621
ɛ:timeT1	0.209	0.605	108.332	0.345	0.731
i:timeT1	**1.247**	**0.597**	**160.9**	**2.089**	**0.038**
ʌ:timeT1	0.712	0.607	112.951	1.172	0.244
ɪ:status2	0.774	0.623	43.312	1.241	0.221
ɛ:status2	**1.704**	**0.489**	**57.095**	**3.485**	**0.001**
i:status2	0.929	0.604	41.264	1.538	0.132
ʌ:status2	0.384	0.627	52.639	0.612	0.543
timeT1:status2	0.066	0.487	148.994	0.135	0.893
ɪ:timeT1:status2	-0.482	0.775	303.965	-0.622	0.535
ɛ:timeT1:status2	-0.22	0.687	160.102	-0.32	0.749
i:timeT1:status2	-0.889	0.787	382.496	-1.13	0.259
ʌ:timeT1:status2	0.283	0.781	255.923	0.362	0.717

Figure 2. Plot of estimated vowel height values for time 1 (left) and time 2 (right), separated by word status ("old" and "new"). Plotted with 95% Confidence Intervals.



The estimated values obtained from the model suggest that, in new words at time 2, the children differentiated between the English /ɛ/ and /æ/ in terms of height. The contrast appears to be much better separated in new words than in old words, with both vowels showing significant change in height: in new words, /æ/ is more open and /ɛ/ is more closed.

The only vowel that showed significant difference between time 1 and 2 was /i/, which was produced as more closed in time 2, moving away from the lower /ɪ/. In the case of /æ/ and /ʌ/, there appears to be some change between time 1 and time 2, irrespective of word status, though this change did not reach significance. In time 2, especially in new words, this change seems to contribute to a better separation of /æ/ and /ʌ/ in height.

Overall, the children seem to better separate contrasting low vowel pairs in new words, rather than in old words, especially the /ɛ-æ/ pair. Productions of new words at time 2 indicate more of a separation of the /æ-ʌ/ contrast. The high vowels do not show an effect of status, and seem to be separated best at time 2, possibly indicating an improvement in time in producing the vowels at a different height.

4.1.2 Vowel retraction

The model for learners' retraction of target vowels was estimated using the formula *F2e-F0e ~ vowel * time * status + (1 + time | speaker) + (1 | word)*. This model's conditional R^2 value is 0.55, with the random effects accounting for a large portion of the variability described by the model (marginal $R^2 = 0.41$).

The p-values for all vowels except /ɛ/ reached significance, indicating that they differed from /æ/ in retraction in new words at time 2. The vowel /ʌ/ was produced as more retracted than /æ/ (β = -1.373, SE = 0.475, t = -2.888, p = 0.006). The vowel /ɛ/ was somewhat more front than /æ/, though the difference did not reach significance (β = 0.647, SE = 0.427, t = 1.516, p = 0.139). The vowels /i, ɪ/ were separated by about 1 ERB at time 2 in new words, indicating some differentiation of the two vowels (see Table 6).

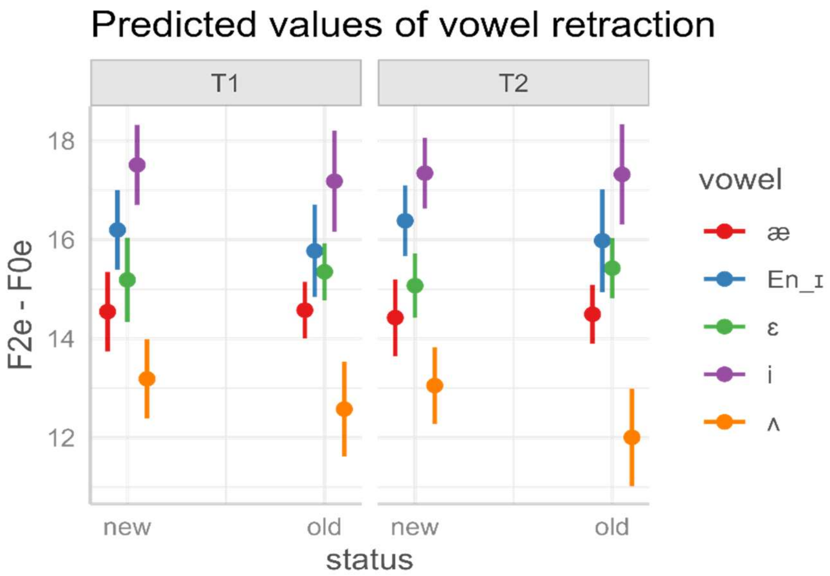
With respect to retraction, no vowel showed a significant change as a function of status or time. However, the vowel /ʌ/ was more retracted in status "old" than in status "new" by 1.115 ERB (β = -1.115, SE = 0.691, t = -1.613, p = 0.116). This is the only indication of a better separation of vowels in old, rather than new words; however, this difference did not reach significance (see Figure 3).

Table 6. The coefficient table for predicted vowel retraction of the five L2 English vowels spoken by the learners.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	**14.425**	**0.395**	**37.801**	**36.513**	**< 0.001**
ɪ	**1.959**	**0.453**	**31.735**	**4.325**	**< 0.001**
ɛ	0.647	0.427	33.921	1.516	0.139
i	**2.922**	**0.454**	**32.696**	**6.437**	**< 0.001**
ʌ	**-1.373**	**0.475**	**36.334**	**-2.888**	**0.006**
timeT1	0.122	0.47	81.844	0.259	0.796
status2	0.069	0.405	35.19	0.17	0.866
ɪ:timeT1	-0.308	0.574	144.218	-0.537	0.592
ɛ:timeT1	-0.006	0.591	130.898	-0.01	0.992
i:timeT1	0.044	0.566	169.467	0.078	0.938
ʌ:timeT1	0.014	0.594	121.202	0.024	0.981

ɪ:status2	-0.471	0.704	28.638	-0.668	0.509
ɛ:status2	0.285	0.531	36.101	0.536	0.595
i:status2	-0.094	0.693	25.042	-0.136	0.893
ʌ:status2	-1.115	0.691	33.064	-1.613	0.116
timeT1:status2	-0.039	0.467	137.298	-0.083	0.934
ɪ:timeT1:status2	0.018	0.705	283.49	0.025	0.98
ɛ:timeT1:status2	-0.15	0.652	175.426	-0.23	0.818
i:timeT1:status2	-0.265	0.705	348.181	-0.376	0.707
ʌ:timeT1:status2	0.473	0.719	228.217	0.658	0.511

Figure 3. Plot of estimated vowel retraction values (in ERB) for time 1 (left) and time 2 (right), separated by word status ("old" and "new"). Plotted with 95% Confidence Intervals.



The results obtained from the models suggest that the learners produced /æ, ʌ/ as different in retraction and height in both old and new words, and at both time 1 and 2. The vowels /ɛ, æ/ seem to have been differentiated only in height, and only in new words (irrespective of time). An inspection of Figure 4 reveals that while /æ/ overlapped less with /ɛ/ in new words, the overlap is still large, and some tokens of /æ/ were /ɛ/-like in both height and retraction. Further, the overlap of /æ, ʌ/ is also large in both hold and new words. The vowel /æ/ is noticeably variable at both time 1 and time 2. The vowel /ʌ/ also seems very variable in old words, and though it is much less variable in new words, it shows great overlap with /æ/. Word status did not affect the production of /i, ɪ/, but /i/ showed a significant change in time, becoming more closed at time 2. Figure 5 reveals that despite this change, the vowels still overlap greatly at both time 1 and time 2.

Figure 4. Low vowels produced by the learners in old and new words. The ellipses use 68% confidence levels.

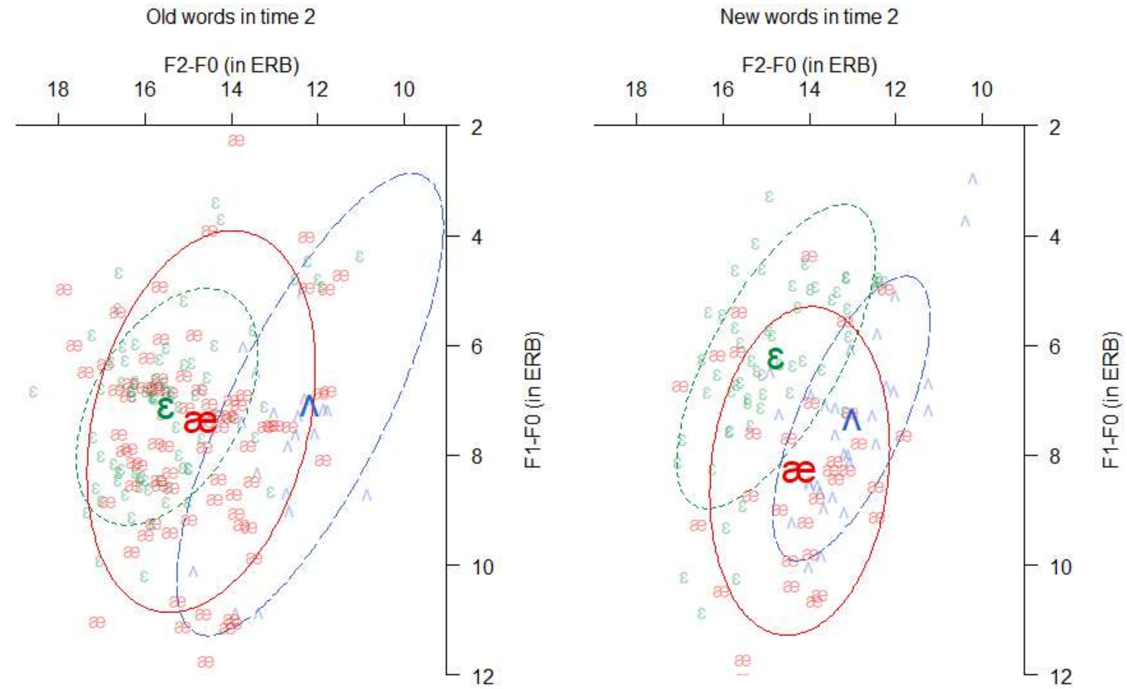
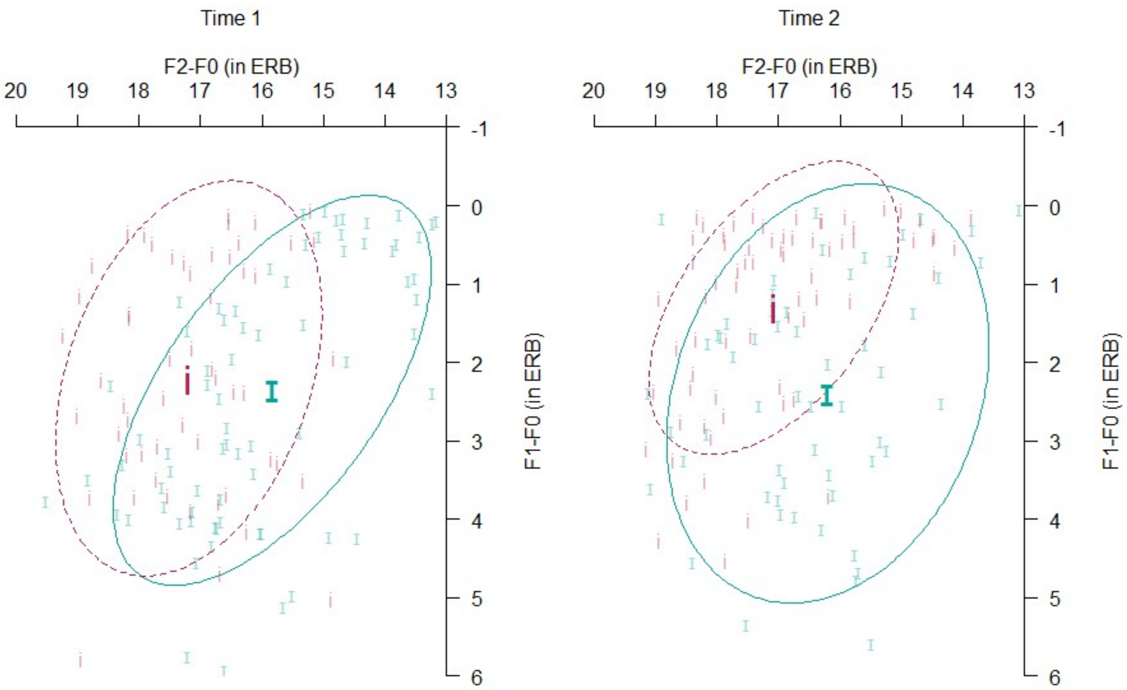


Figure 5. High vowels produced by the learners in time 1 and 2. The ellipses use 68% confidence levels.



4.1.3 Individual differences

Individual speakers' deviations from the grand intercept of the height model (/æ/ in time 2 in new words in the intercept in Table 5) were examined. Speakers who deviated most from the estimated height value for /æ/ were speaker 1, whose /æ/ productions were more open (by 1.02 ERB), and speaker 8, whose productions were more closed (by 0.88 ERB). For speaker 8, /æ, ɛ/ seem to be merged, with /æ/ being more variable in height. It should be noted that she was a speaker of L1 Slovak, though she used (and was exposed to) Czech as her dominant language for most of her life. Input in Slovak may have influenced her English vowel realisations. Speaker 1 shows a tendency in the opposite direction, overlapping /æ/ more with /ʌ/, and very little with /ɛ/ (Figure 6).

In retraction, a great deviation from the grand intercept (/æ/ in time 2 in new words the intercept in Table 6) was found in speaker 10 (by 0.87 ERB), who produced /æ/ as more back than those modelled for the entire group. More retracted productions of /æ/ suggest perceived similarity between tokens of /æ/ and the L1 vowel /a/. However, the /æ/ realisations of speaker 10 seem to be variable enough to suggest confusion caused by the similarity of /æ/ to both L1 vowels /ɛ/ and /a/.

Figure 7. English low vowels of speaker 1 (left), who produced /æ/ as more open than the other learners, and s8 (right), who produced /æ/ as more closed than the other learners. The plots show averages and tokens from both times. Ellipses use 68% confidence levels.

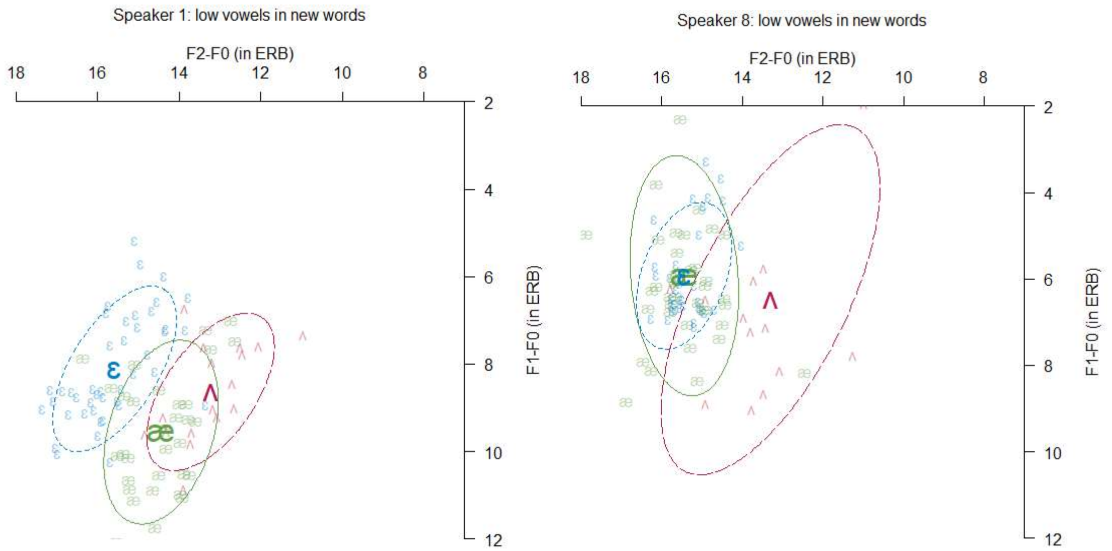
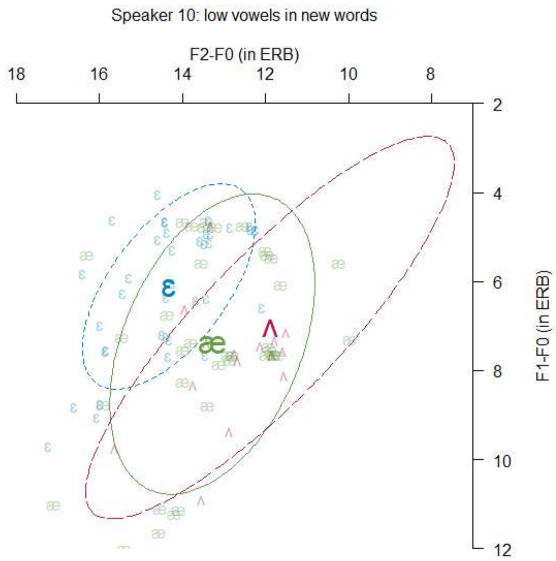


Figure 8. Vowel plot of productions of speaker 10, whose /æ/ was more retracted than that of the other learners. The vowel tokens are taken from new word realisations across time. Ellipses use 68% Confidence levels.



4.2 Czech vowels produced by the learners

To estimate height and retraction of the Czech front vowels produced by the learners in time 1 and time 2, two linear mixed-effects models were used to model height and retraction as a function of the fixed effects *vowel* and *time*, with an interaction between the two. The random effects were *time*, with by-speaker varying slopes, and *word*, with varying intercepts. Time was treatment-coded with time 1 as the reference level. Only seven learners were included due to lack of data from speaker 2 and speaker 10. The data supplied to the model were 202 word tokens (19 types) produced by the seven learners. The Intercept in both models was set to the vowel /ɛ/.

4.2.1 Vowel height

The model for learners' height of target vowels was estimated using the formula *F1e-F0e ~ vowel * time + (1+time | speaker) + (1 | word)*. The total explanatory power of the model was substantial (conditional $R^2 = 0.85$; marginal $R^2 = 0.83$).

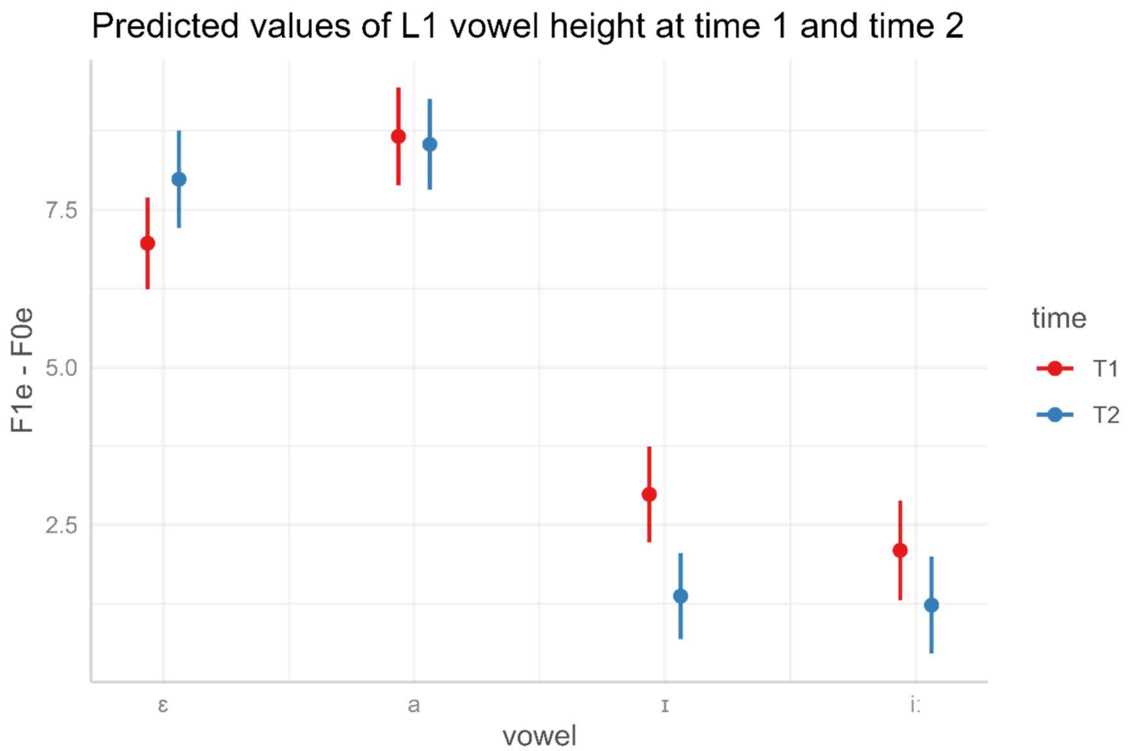
There was a significant effect for the vowels /a, ɪ, iː/ relative to the vowel /ɛ/ in the Intercept, meaning that productions of all these vowels differ significantly from productions of /ɛ/ (see Table 7).

Some vowel-time interactions reached significance, indicating that the learners' L1 vowels differed substantially between time 1 and 2. There was a shift in the vowel /ɪ/, which became higher at time 2 ($\beta = -2.63$, $SE = 0.671$, $t = -3.917$, $p = 0.002$, and /i/, which was also higher at time 2 ($\beta = -1.885$, $SE = 0.706$, $t = -2.669$, $p = 0.021$). The vowel /ɛ/ was also higher at time 2, though this change did not reach significance ($\beta = 1.016$, $SE = 0.492$, $t = 2.064$, $p = 0.06$). The changes in the two high vowels made them more similar to each other in height at time 2 than they were at time 1. So much so that at time 2, the high vowels did not seem to be different in height. The vowel /a/ was the only one that did not change in height between time 1 and 2.

Table 7. The coefficient table for predicted vowel height of Czech vowels spoken by the learners in time 1 and 2.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	**6.964**	**0.37**	**16.145**	**18.825**	**0**
a	**1.695**	**0.461**	**10.147**	**3.675**	**0.004**
Cz_ɪ	**-3.98**	**0.454**	**10.928**	**-8.775**	**0**
iː	**-4.869**	**0.468**	**11.563**	**-10.394**	**0**
timeT2	1.016	0.492	12.436	2.064	0.06
a :timeT2	-1.142	0.686	9.981	-1.665	0.127
Cz_ɪ :timeT2	**-2.63**	**0.671**	**11.179**	**-3.917**	**0.002**
iː :timeT2	**-1.885**	**0.706**	**11.553**	**-2.669**	**0.021**

Figure 9. Plot of estimated vowel height values for the learners' Czech vowels in time 1 and 2.



4.2.2 Vowel retraction

The model for learners' retraction of target vowels was estimated using the formula *F2e-F0e ~ vowel*time + (1+time | speaker) + (1 | word)*. The marginal R^2 value for this model was 0.69 and the conditional R^2 value was 0.78.

There was again a significant effect for the vowels /a, ɪ, iː/ relative to the vowel /ɛ/ in the Intercept, meaning that all these vowels differ significantly in retraction from /ɛ/ at time 1. The time-vowel interaction reached significance only for the vowel /ɛ/, which was produced as more back at time 2 ($\beta = -1.024$, SE = 0.391, $t = -2.622$, $p = 0.019$) (see

Figure **10**). The vowel /a/ also became more retracted at time 2, though this change did not reach significance, as /a/ remained similarly different from the reference level /ɛ/ at both times, because /ɛ/ shifted together with /a/ (see

Table **8**). Both /ɛ/ and /a/ became more retracted at time 2, with the distance between them staying approximately the same. The vowel /iː/, which did not seem to be separated in height from /ɪ/, appears to be produced as different when it comes to retraction, with /iː/ being more front than /ɪ/ (by about 1.3 ERB at time 1), and with /ɪ/ becoming even more front in time 2 (see Figure 11).

Table 8. The coefficient table for predicted vowel height of Czech vowels spoken by the learners at time 1 and 2.

	Estimate	**Std. Error**	**df**	**t value**	**Pr(>|t|)**
(Intercept)	**15.677**	**0.354**	**11.006**	**44.328**	**0**
a	**-1.711**	**0.322**	**10.228**	**-5.312**	**0**
Cz_ɪ	**0.974**	**0.317**	**10.882**	**3.068**	**0.011**
iː	**2.363**	**0.328**	**11.519**	**7.211**	**0**
timeT2	**-1.024**	**0.391**	**14.776**	**-2.622**	**0.019**
a :timeT2	-0.123	0.479	10.153	-0.258	0.802
Cz_ɪ :timeT2	0.605	0.47	11.31	1.286	0.224
iː :timeT2	0.972	0.495	11.855	1.964	0.073

Figure 10. Plot of estimated vowel retraction values of the learners' Czech vowels in time 1 and 2 (in ERB). Plotted with 95% Confidence Intervals.

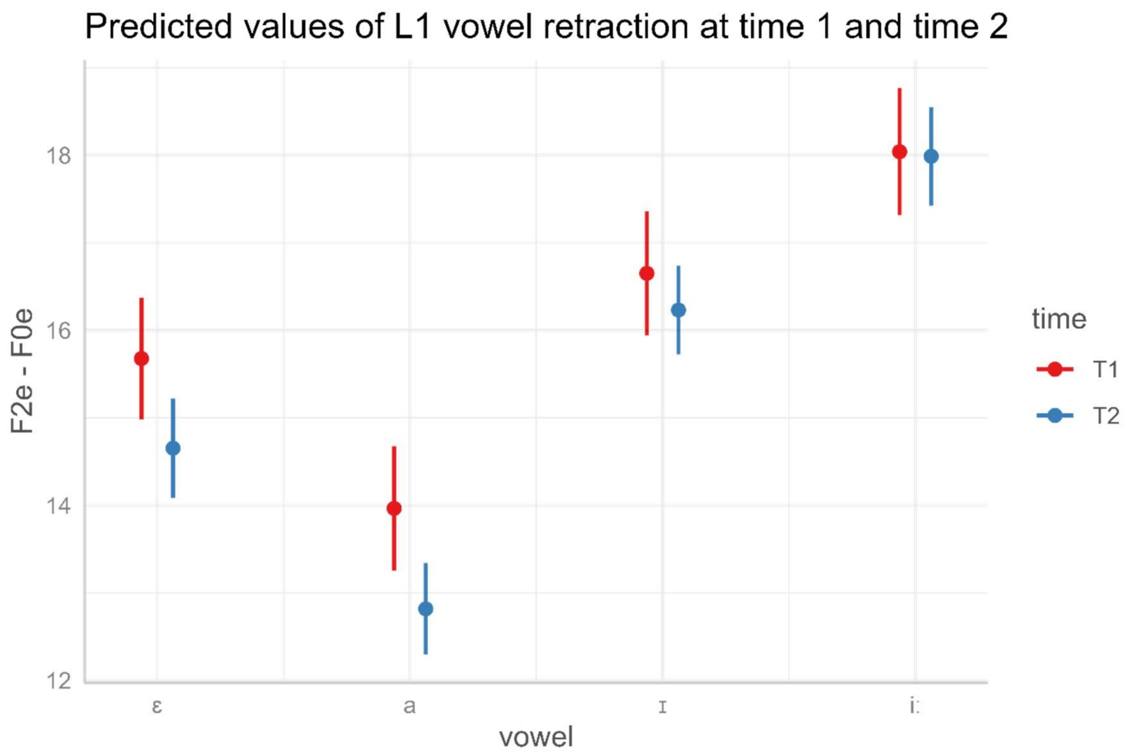
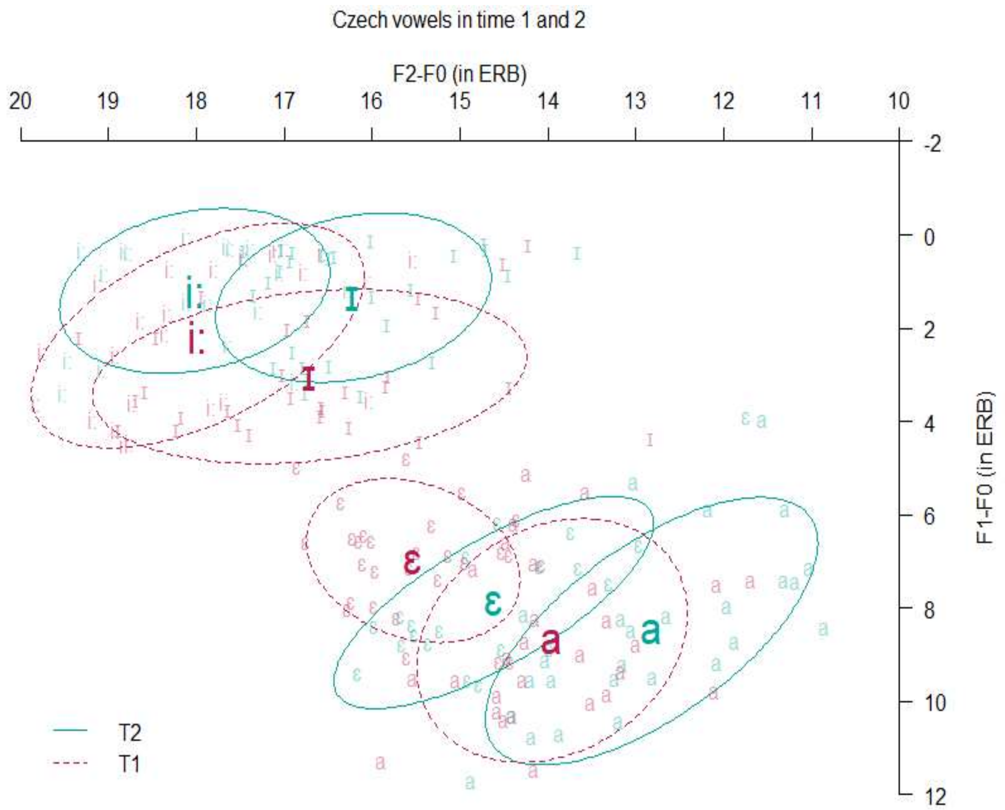


Figure 11: Czech vowels in time 1 (blue) and time 2 (red) produced by the seven learners. Ellipses use 68% confidence intervals.



4.2.3 The learners compared with Czech monolinguals

Figure 12 illustrates the shifts in the seven learners' Czech vowels when compared to Czech vowels produced by Czech monolinguals. It seems that at time 1, especially the vowels /ɪ, ɛ/ produced by the learners were more similar to produced by the monolinguals. The raising of the high vowels, and lowering and retracting of /ɛ/ can thus also be seen when comparing the learners' time 2 production with the monolinguals' realisation of the vowels.

Figure **13** shows the learner's realisations in new words at time 2 compared to data gathered from YouTube videos with SSBE monolinguals. The monolinguals' realisations appear very variable, which may be due to the lower quality of the audio, or the monolinguals coming from a more versatile dialect background. Nevertheless, the learners' realisations appear to approximate those of the monolinguals, especially when it comes to the low vowels.

Figure 12: Czech vowels produced by the learners at time 1 (left) and time 2 (right) compared to vowels produced by Czech monolinguals. Ellipses use 68% confidence levels.

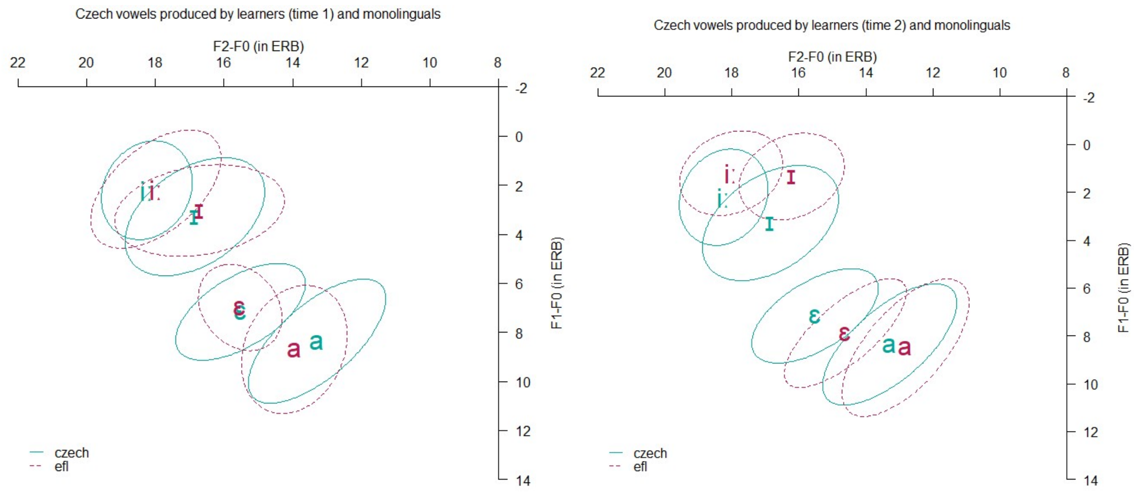
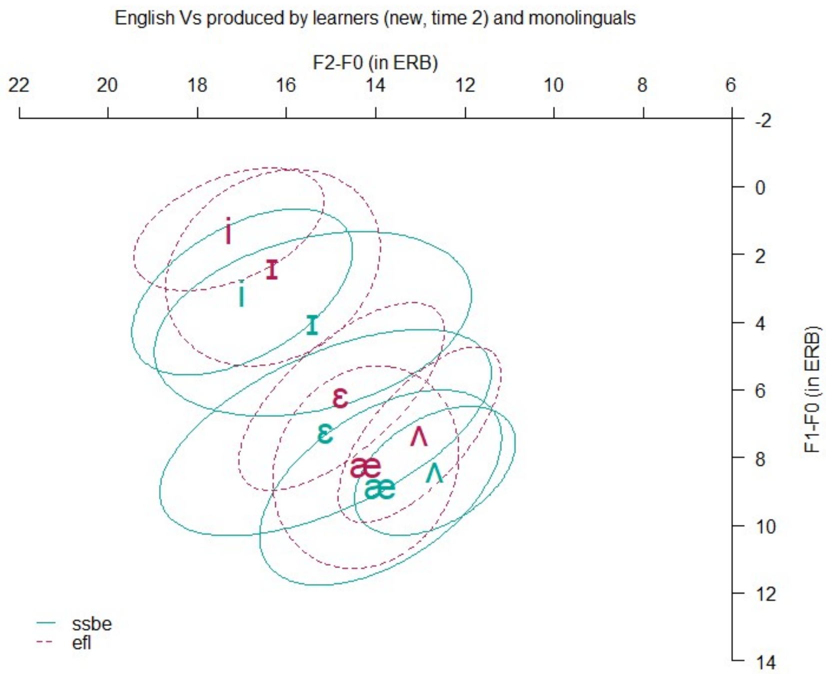


Figure 13. English vowels produced by the learners (in new words at time 2) and by SSBE monolinguals from YouTube videos. Ellipses use 68% confidence levels.



4.3 Perception of L2 vowels in words and non-words

The data from the perception tests were plotted, separating the nine learners into two groups. Four learners tended to accept most words produced by the teacher (Group 2, Figure 15), while the remaining five had an accept rate that indicated better discrimination of L2 vowels (Group 1,
Figure **14**).

The more successful learners showed similar accept rates for all vowel replacements, rejecting non-words where /æ/ was replaced with /ɛ/ and /i/ with /ɪ/ slightly more often than non-words pronounced with /æ/ and /i/. They rejected /æ/ replacements in

63% of tokens. Mispronunciations with /ɛ/ had the highest reject rate, the learners in group 1 rejected 75% of them. The lowest reject rate was 60%, for mispronunciations with /ɪ/. As for correct productions, 100% of correct /ɛ/ and /i/ tokens were accepted, 91% of /æ/ tokens were accepted, and the vowel /ɪ/ and filler /ɑ/ showed the lowest accept rate: 80%. Surprisingly, speaker 8 was among the learners who rejected the words more successfully, even though she did not seem to differentiate English /ɛ/ and /æ/ in production (see

Figure **7**).

The less successful group also accepted all correct /i/ words and accepted all other vowels in more than 85% of tokens. However, in the words with replaced vowels, only 19% of words with /æ/ and /ɛ/ replacements was rejected, and less than 13% of tokens with /i, ɪ/ replacements were rejected.

Figure 14. Percentage of “accept” (green) and “reject” (red) reactions in correctly produced words (left) and words produced with a replaced vowel (right) for the five children who performed better in the perception task.

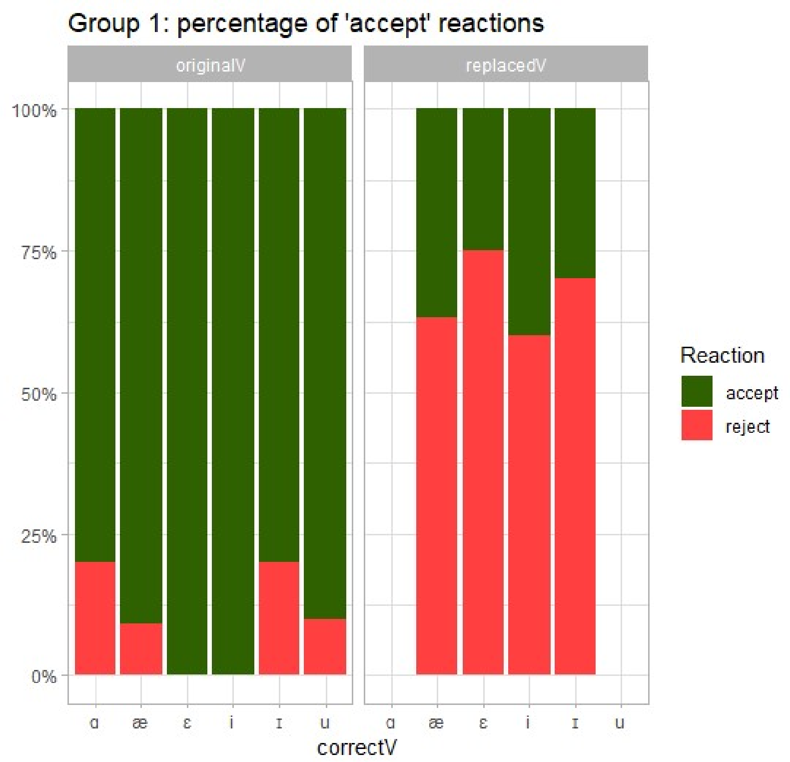
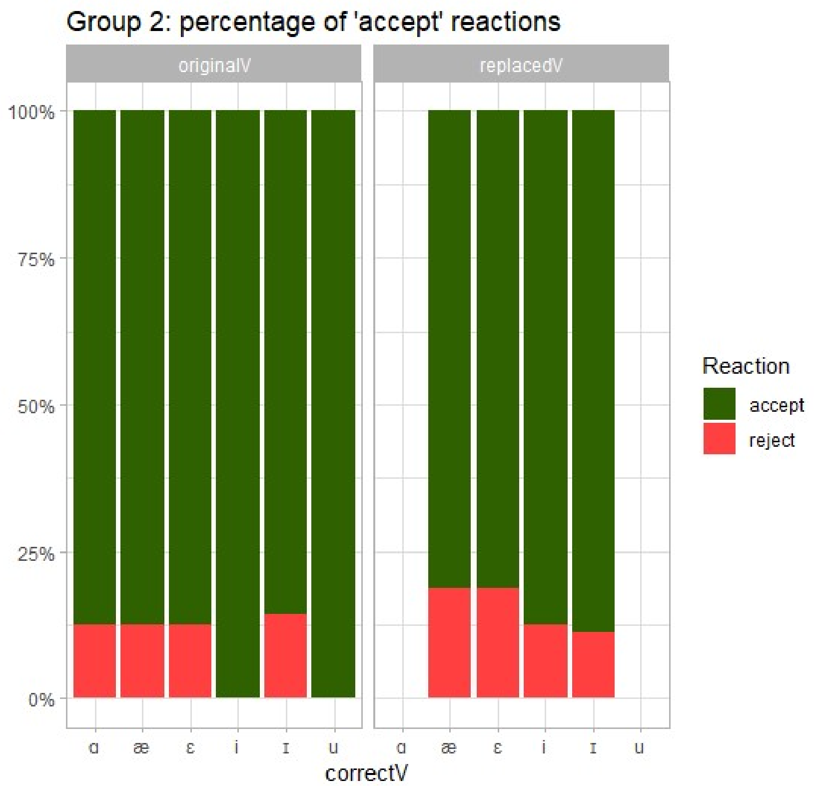


Figure 15. Percentage of “accept” (green) and “reject” (red) reactions in correctly produced words (left) and words produced with a replaced vowel (right) for the four learners who tended to accept any production.



5 Discussion

This study examined production and perception of the L2 vowels /i, ɪ, ɛ, æ, ʌ/ in preschool EFL learners throughout three months, and investigated drift in their L1 vowels. The first aim was to address whether child EFL learners differentiate contrasting L2 vowels in production, considering the effect of source of input and time. This effect was addressed by comparing acoustic measurements of vowel tokens produced in old words (heard primarily from speakers who did not preserve the /æ-ɛ/ contrast) and new words (heard from a speaker that preserved L2 contrasts). The effect of time on L2 vowel production was assessed by comparing vowel realisations from the first six weeks of the experiment with those from the last six weeks. Additionally, a task based on L2 vowel substitutions was conducted to assess the success with which the learners rejected non-word pronunciations. The second aim was to examine whether L2 learning impacts the learners' production of L1 vowels by comparing vowel production in two sessions separated by 10 weeks.

5.1 Production of L2 vowels

5.1.1 The vowels /ɛ, æ, ʌ/

The L2 vowels /ɛ, æ, ʌ/ were selected because they represent different L1-L2 category mappings based on input source. Based on Pierrehumbert (2001), it was suggested that the degree to which the learners produce the vowels contrastively may depend on the source of input which provided the word exemplars saved by the children, and those exemplars' activation strength. As predicted, the vowel pair /ɛ, æ/ did not appear to be differentiated in old words, in which both /ɛ, æ/ were argued to be represented by /ɛ/-like vowel exemplars. In words provided by the input source which preserved the contrast, the learners produced /æ, ɛ/ as better separated in height, however, much overlap was still observed, and the degree of separation of the vowels also varied in individual learners. Similarly, separation of /æ-ʌ/ was predicted to be influenced by input source, with the two vowels being more similar to each other in new words under the influence of /æ/ being lower, and less similar in old words. Though the vowels appeared to be sufficiently separated in both old and new words, a large overlap was observed in both.

Overall, the learners separated the vowels /ɛ, æ, ʌ/ better in new words, but the variability of /æ/ caused an overlap with both /ɛ, ʌ/. This indicates a persisting struggle to separate /æ/ from the /ɛ, ʌ/ and suggests an underdeveloped L1-L2 category for /æ/, which is likely associated with /ɛ/ under the influence of non-native input, and which is being newly associated with /ʌ/ under the influence of input that separates /ɛ, æ/. Additionally, even in new words, some tokens of /æ/ were produced as higher and more front, and so more /ɛ/-like. Inspection of individual vowel plots revealed that even the learner who showed the least overlap between /æ, ɛ/ in new words, still produced some /ɛ/-like instances of /æ/. This further supports the existence of a common phonological category for /æ/ and /ɛ/. The observed overlap of /æ, ɛ/ is consistent with research on young adult Czech-English bilinguals (Šimáčková and Podlipský 2018), who also overlapped the two vowels. Like the low proficiency child learners in Yang and Fox (2017), the learners' L2 vowels showed much overlap (despite the source of input effect), indicating that the learners have not yet established separate categories for L2 vowels, but they were learning to produce the vowels as different in quality.

An effect of time on learning to separate /ɛ, æ/ in old words was predicted to be possible, but unlikely. The impact of word status on the production of /ɛ, æ/ can be accounted for by a preference for commonly heard variants (Pierrehumbert 2001). In the old words, the most commonly heard variants of /æ/ were /ɛ/-like, leading the children to produce the words accordingly. This may also explain why there was no effect of time for old words. There were still more /ɛ/-like exemplars of /æ/ associated with old words, which were also being reinforced through non-native input outside lessons, making these exemplars more strongly activated, as opposed to the less often heard tokens containing /a/-like exemplars of /æ/ provided during lessons.

A question arises from the fact that the learners did not seem to extend the more separated production of /æ, ɛ/ to old words. Did they realise that the old and new words contain the same vowel, or did they take old words as containing a different (more /ɛ/-like sound) than the new words? Splitting the /æ/ sounds in this way would indicate lexical misrepresentation, which has been reported even for highly proficient Czech EFL learners (Šimáčková and Podlipský 2018).

When it comes to assessing the effect of input source on the learners' vowels realisations, the study could be improved by conducting an analysis the vowel realisations heard by the children outside class, so that the influence of this input could be evaluated, and not only assumed.

5.1.2 The vowels /i, ɪ/

The vowel pair /i, ɪ/ was predicted to be differentiated in quality depending on the separation of the Moravian Czech /iː, ɪ/, and learners were assumed to reuse the L1 quantity cue as the primary means to contrast the L2 pair. For /i, ɪ/, the effect of word status or time was predicted to be unlikely, as the L2 vowels were expected to map onto L1 categories based on duration in a two-category assimilation, and unlike /æ/, neither of the closed vowels was expected to vary in mapping according to input source. As predicted, input source did not seem to affect production of the two vowels.

Contrary to the prediction, the results showed that /i/ became higher at time 2, indicating better separation of the vowels. However, even though there seems to be some improvement, much overlap was still observed, suggesting that despite the raising of /i/, the qualitative contrast was not preserved sufficiently in production at neither time 1 nor time 2. The L2 vowels /i, ɪ/ thus, like /æ/, indicated that the learners had not yet established L2 categories contrasted in quality, and the observed raising of /i/ did not seem to contribute to a better separation of the vowels. This is taken to indicate that the learners did reuse the L1 quantity cue, which Moravian Czech speakers rely on in their native language (Podlipský, Skarnitzl and Volín 2009), to contrast the vowels at both time 1 and time 2, overlooking the quality contrast.

The large overlap between /i, ɪ/ stands in contrast to findings of Šimáčková and Podlipský (2018) who reported that proficient Czech bilinguals produced the vowels as different in both quality and quantity. This underscores the importance of the nature of learners' L1 categories as a possible factor influencing L2 vowel acquisition (Chládková and Podlipský 2011), as the learners in Šimáčková and Podlipský (2018) may have come from a different dialectal background. However, it should also be noted that they, as proficient bilinguals, represented an entirely different population from the learners studied here, so the contrary finding may be a product of differing amounts of L2 experience, rather than differences in L1 categories.

5.2 Phonetic drift in L1 vowels

Seven of the nine learners produced Czech words containing the target vowels /iː, ɪ, ɛ, a/ in two sessions separated by 10 weeks, one session took place in the first month of the experiment, the other in the third. The aim was to investigate whether any change in L1 vowel production had occurred while the children had been learning the L2. Phonetic drift in the L1 was expected based on findings of Chang (2012) and especially of Yusa et al. (2010), who found L1 drift in preschool EFL learners. Like in these studies, a drift effect was observed, specifically, changes in the production of /ɪ, ɛ, a/. The vowel /ɪ/ raised, so /iː, ɪ/ came to be differentiated in retraction, rather than height; /ɛ/ lowered, and /a/ lowered and retracted, moving away from the lowered /ɛ/.

The drift effect found could be due to a variety of reasons. It must be noted that vowel systems of individuals, including monolinguals, are dynamic, changing throughout life. The children's L1 vowels are still developing (Lee, Potamianos and Narayanan 1999, Walley and Flege 1999), and changes in their vowel inventory may also be caused by developmental factors. However, the reasons behind phonetic drift may also attributed to the learners' need to enhance perceptual distinctiveness between the vowels in the learner's L1-L2 system (Guion 2003), which may have also played a role in the L1 vowel changes observed in this study. Also, the learners' realisations of high vowels at time 2 are similar to the teachers', so the change may also be due to the learners imitating the teacher's pronunciation.

However, to address the nature of the observed phonetic drift, and directly assess the interaction between the learners' L1 and L2 categories, more data would be needed, which could be used to model both Czech and English vowels, allowing for a cross-language comparison between the vowels, and an assessment of differentiation of similar L1 and L2 vowels. This was not done here as conducting tests to predict production of both Czech and English vowels for the seven learners would decrease the statistical power of the study. Unfortunately, as the models for Czech vowels only use data from seven of the learners, direct comparison of the results dealing with L2 contrast acquisition and L1 drift is not possible. Hence, the nature of the drift effect (assimilation vs. dissimilation) cannot be confidently determined.

5.3 Vowel perception in (non-)words

The learners were asked to judge the acceptability of the teacher's single-word productions featuring either the correct vowel, or a replaced vowel. Predictions for vowel perception mirrored those for production of old words: both /æ-ɛ/ and /i-ɪ/ were expected to be discriminated using duration. Five children performed at above-chance levels in rejecting the non-words, whereas four children accepted the majority of both words and non-words. The more successful learners showed similar success rates for all vowel replacements.

The reasons for why four learners seem to have performed worse than the remaining five are unclear. There are many factors that may have affected the learners' performance, including task design and individual variation in the degree of precision of the learners' L2 perceptual categories. The less successful learners probably did not attend to vowel duration, as the consonantal context was not controlled, so duration is assumed to have been variable. Highlighting coevolution of production and perception (Flege and Bohn 2021), one of the learners performed well on the perception task for /æ, ɛ/, though she seemed to merge them in production.

L2 learners have been shown to perform differently in tasks that involve processing at the lexical versus the phonetic level (Hazan 2007). As the task involved lexical processing, knowledge of the words was necessary to recognise a vowel substitution. Also, even though a puppet was used in both the in-class perception trials and during the perception test, as the teacher was producing the words, possibly making some of the learners more inclined to accept her productions. In the future, assessing children's perception of isolated L2 vowels as well as vowels in lexical context may help explain individual variability.

The perception of learners who showed better accept rates still appears to be imprecise, indicating wide perceptual categories, similarly to the L1 listeners in Walley and Flege (1999). The learners with high accept rates may have considered the substituted vowels atypical instances of the vowel categories, but still acceptable, which would signal a variable perceptual category. This is in accordance with the findings of Simon, Sjerps and Fikkert (2014), who found that child L2 learners accepted many incorrect vowel pronunciations in words, indicating underdeveloped phonological representations of L2 vowels.

6 Conclusion

Young children learning EFL are an understudied population that typically receives less native speaker input than SLA learners. This study highlights the influence of non-native input on the development of L2 vowels in young EFL learners. It shows that the effect of different sources of non-native input on L2 vowel production can be accounted for by referring to the L1-L2 vowel mapping pattern learners establish separately for differing sources of input, and by learners' preference for commonly heard variants of words, assuming that word exemplars in a learners' mind also include phonetic information. If learners map L2 sounds onto different L1 categories based on different sources of input, they may split their vowel realisations in words according to the input source which provided the most activated exemplars of those words. If L2 sounds are consistently mapped onto L1 categories, they will be realised more consistently across word productions. These findings underscore the importance of how different sources of input realise L2 sounds, as their realisation affects how learners produce those L2 sounds, with the input possibly contributing to, or inhibiting L2 sound category development. This study further observed change in the learners' L1 vowels during a 10 week period of learning, suggesting that phonetic drift also occurs in inexperienced child EFL learners. Additionally, the observed individual variability in accept rates of word mispronunciations based on L2 vowel substitutions underlines the many factors that may affect preschoolers' performance in perception tasks.

7 References

Asher, J. J. 1969. “The total physical response approach to second language learning.” *Modern Language Journal* 53, no. 1: 3–17. https://doi.org/10.2307/322091.

Baker, Wendy, Pavel Trofimovich, James E. Flege, Molly Mack, and Randall Halter. 2008. “Child—adult differences in second-language phonological learning: The role of cross-language similarity.” *Language and Speech* 51, no. 4: 317–342.

Bates, Douglas, Martin Maechler, Ben Bolker, Steve Walker. 2015. “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software* 67, no. 1: 1–48. https://doi.org/10.18637/jss.v067.i01.

Best, Catherine T., M. Tyler, O. Bohn. 2007. “Nonnative and second-language speech perception.” In *Language experience in second language speech learning*, edited by M. Munro, 13–34. https://doi.org/10.1075/lllt.17.

Bjelaković, A. 2016. “The vowels of contemporary RP: Vowel formant measurements for BBC newsreaders.” *English Language & Linguistics* 21, no. 3: 501-532. https://doi.org/10.1017/S1360674316000253.

Boersma, Paul and David Weenink. 2022. “Praat: doing phonetics by computer” [Computer program]. Version 6.2.23, retrieved 8 October 2022 from http://www.praat.org/.

Bosch, Laura and Marta Ramon-Casas. 2011. “Variability in vowel production by bilingual speakers: Can input properties hinder the early stabilization of contrastive categories*?” Journal of Phonetics* 39: 514-526.

Bundgaard-Nielsen, Rikke L., Catherine T. Best, and Michael D. Tyler. 2011. “Vocabulary size is associated with second-language vowel perception performance in adult learners.” *Studies in second language acquisition* 33, no. 3: 433–461.

Chang, Charles B. 2012. “Rapid and multifaceted effects of second-language learning on first-language speech production.” *Journal of Phonetics* 40: 249–268.

Chládková, Kateřina, Václav J. Podlipský. 2011. “Native dialect matters: perceptual assimilation of Dutch vowels by Czech listeners.” *The Journal of the Acoustical Society of America* 130, no. 4: EL186–EL192. https://doi.org/10.1121/1.3629135.

Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Cambridge: Cambridge University Press.

Cruttenden, Alan. 2014. *Gimson's pronunciation of English.* London: Routledge.

Daniel R. McCloy. 2016. “phonR: tools for phoneticians and phonologists.” R package version 1.0-7. http://drammock.github.io/phonR/.

De Haro, Alfredo Herrero. 2022. “praat-script.-f1-f2-duration-pitch-intensity” [Praat script]. Retrieved from https://aherrerodeharo.com/research-resources/.

Derdemezis, Ekaterini, Houri K. Vorperian, Ray D. Kent, Marios Fourakis, Emily L. Reinicke, and Daniel M. Bolt. 2016. “Optimizing Vowel Formant Measurements in Four Acoustic Analysis Systems for Diverse Speaker Groups.” *American journal of speech-language pathology* 25, no. 3: 335–54. https://doi.org/10.1044/2015_AJSLP-15-0020.

Dörnyei, Zoltán. 2009. “The 2010s Communicative language teaching in the 21st century: The ‘principled communicative approach.’” *Perspectives* 37, no. 2: 33–43.

Doron, Helen. 1993. “Teaching English to Infants and Primary Schoolchildren in Israel--A Special Project.” E*ducational Media International* 30, no. 3: 119–123. https://doi.org/10.1080/0952398930300302.

Doron, Helen. 2011. *The Music of Language.* Kindle Edition. https://www.amazon.com/Music-Language-Helen-Doron-ebook/dp/B0062YQEJ.

Escudero, Paola. 2001. “The role of the input in the development of L1 and L2 sound contrasts: language-specific cue weighting for vowels.” In Proceedings of the 25th annual Boston University conference on language development, vol. 1, pp. 250-261. Somerville, MA: Cascadilla Press.

Flege, J. E. 1987. “The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification.” *Journal of Phonetics* 15: 47–65.

Flege, J. E. 1995. “Second language speech learning: Theory, findings and problems.” In *Speech perception and linguistic experience: Issues in cross-language research*, edited by W. Strange, 233–277. Baltimore: York Press.

Flege, J., and O. Bohn. 2021. “The Revised Speech Learning Model (SLM-r).” In *Second Language Speech Learning: Theoretical and Empirical Progress,* edited by R. Wayland, 3–83. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108886901.002.

Flege, James E., and Wieke Eefting. 1987. “Production and perception of English stops by native Spanish speakers.” *Journal of Phonetics* 15: 67–83.

Flege, James E., Carlo Schirru, and Ian R. A. MacKay. 2003. “Interaction between the native and second language phonetic subsystems*.” Speech communication* 40, no. 4: 467–491.

Glasberg, B. R., and B. C. J. Moore. 1990. Derivation of auditory filter shapes from notched-noise data. *Hearing Research* 47, no. 1-2: 103–138. https://doi.org/10.1016/0378-5955(90)90170-T.

Guion, Susan G. 2003. “The Vowel Systems of Quichua-Spanish Bilinguals.” *Phonetica* 60: 98–128.

Hazan, Valerie. 2007. “Second Language Acquisition and Exemplar Theory.” ICPhS XVI.

Helen Doron English. 2022. “Our Courses” [website]. Accessed December 7, 2022. https://helendoron.com/english/our-courses/.

Kartushina, Natalia, Alexis Hervais-Adelman, Ulrich Hans Frauenfelder, and Narly Golestani. 2016. “Mutual influences between native and non-native vowels in production: Evidence from short-term visual articulatory feedback training.” *Journal of Phonetics* 57: 21–39.

Klimova, B. 2013. “Teaching English to Pre-school Children.” *Procedia - Social and Behavioral Sciences* 93. https://doi.org/10.1016/j.sbspro.2013.09.228.

Ladefoged, Peter, and Keith Johnson. 2014. *A Course in Phonetics*. Stamford: Cengage learning.

Lee, S., Potamianos, A., and S. Narayanan. 1999. “Acoustics of children’s speech: Developmental changes of temporal and spectral parameters.” *The Journal of the Acoustical Society of America* 105: 1455–1468.

Lenth, R. 2023. “emmeans: Estimated Marginal Means, aka Least-Squares Means.” R package version 1.8.4-1. https://CRAN.R-project.org/package=emmeans.

Levy, Helena, and Adriana Hanulíková. 2019. “Variation in children’s vowel production: Effects of language exposure and lexical frequency.” *Laboratory Phonology* 10, no. 1: 1–26. https://doi.org/10.5334/labphon.131.

Machač, Pavel, and Radek Skarnitzl. 2009. *Principles of phonetic segmentation*. Praha: Epocha.

Marian, Blumenfeld, and Kaushanskaya. 2007. “The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals.” *Journal of Speech Language and Hearing Research* 50, no. 4: 940–967. Translated by B. Keřková, M. Večeřová, and K. Knížková (2021), Národní ústav duševního zdraví (National Institute of Mental Health, NIMH-CZ).

Marian, Blumenfeld, and Kaushanskaya. 2007. “The Language Experience and Proficiency Questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals.” *Journal of Speech Language and Hearing Research* 50, no. 4, 940–967. Adapted for use with children by Sirada Rochanavibhata, Margarita Kaushanskaya, and Viorica Marian, Northwestern University.

Morrison, G. S. 2009. “L1-Spanish Speakers’ Acquisition of the English /i/—/ɪ/ Contrast II: Perception of Vowel Inherent Spectral Change.” *Language and Speech* 52, no. 4: 437–462. https://doi.org/10.1177/0023830909336583.

No author. “Soubor orientačních testů pro věkovou skupinu 5 – 6 roků.” https://is.muni.cz/th/tlfx3/soubor_4-5roku.pdf.

Pierrehumbert, J. B. 2001. “Exemplar dynamics: Word frequency, lenition and contrast.” In *Frequency and the emergence of linguistic structure*, edited by J. Bybee and P. Hopper, 137–157. Amsterdam: Benjamins.

Podlipský, Václav J., Radek Skarnitzl, and Jan Volín. 2009. “High front vowels in Czech: A contrast in quantity or quality?” Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH: 132–135. 10.21437/Interspeech.2009-50.

R Core Team. 2023. “R: A language and environment for statistical computing.” R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Šimáčková, Šárka, and Václav J. Podlipský. 2018. Production Accuracy of L2 Vowels: Phonological Parsimony and Phonetic Flexibility. *Research in Language* 16: 169–191. 10.2478/rela-2018-0009.

Šimáčková, Šárka, Václav Jonáš Podlipský, and Kateřina Chládková. 2012. “Czech Spoken in Bohemia and Moravia*.” Journal of the International Phonetic Association* 42, no. 2: 225–32. https://doi.org/10.1017/S0025100312000102.

Simon, Ellen, Matthias J. Sjerps, and Paula Fikkert. 2014. “Phonological representations in children’s native and non-native lexicon.” *Bilingualism: Language and Cognition* 17, no. 1: 3–21.

Skarnitzl, Radek and Jan Volín. 2012. “Referenční hodnoty vokalických formantů pro mladé dospělé mluvčí standardní češtiny” [Reference Values of Vowel Formants for Young Adult Speakers of Standard Czech]. *Akustické listy* 18: 7-11.

Skarnitzl, Radek, and Jana Rumlová. 2019. "Phonetic aspects of strongly-accented Czech speakers of English." *AUC PHILOLOGICA 2019*, no. 2: 109–128.

Skarnitzl, Radek, Pavel Šturm, and Jan Volín. 2016. *Zvuková báze řečové komunikace: Fonetický a fonologický popis řeči* [Aural basis of speech communication: phonetic and phonological description of speech]. Charles University in Prague: Karolinum Press.

Šturm, Pavel, and Radek Skarnitzl. 2011. "The Open Front Vowel /æ/ in the Production and Perception of Czech Students of English." Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, 1161–1164. 10.21437/Interspeech.2011-344.

Tefelnerová, Petra. 2013. "Teaching English in kindergartens: from a survey of existing approaches to designing a convenient methodology." Bachelor thesis, Masarykova univerzita, Pedagogická fakulta.

Tyler, Michael. 2019. "PAM-L2 and phonological category acquisition in the foreign language classroom." In *A Sound Approach to Language Matters – In Honor of Ocke-Schwen Bohn*, edited by Nyvad, Anne Mette, Michael Hejna, Anders Højen, Anna Jespersen and Mette Hjortshøj Sørensen, 607–630. Dept. of English, School of Communication & Culture, Aarhus University. https://doi.org/10.7146/aul.322.218.

Vidnerová, Petra. 2014. "Rozvoj slovní zásoby dětí předškolního věku." Diploma thesis. https://theses.cz/id/vuv3np/11441065.

Viegas, Flávia, Danieli Viegas, and Heidi Elisabeth Baeck. 2015. "Frequency measurement of vowel formants produced by Brazilian children aged between 4 and 8 years." *Journal of voice* 29, no. 3: 292–298. https://doi.org/10.1016/j.jvoice.2014.08.001.

Walley, Amanda C., and James E. Flege. 1999. "Effect of lexical status on children's and adults' perception of native and non-native vowels." *Journal of Phonetics* 27, no. 3: 307–332.

Wang, Yizhou, Bundgaard-Nielsen, Rikke L., Baker, Brett J. and Maxwell, Olga. 2023. "Difficulties in decoupling articulatory gestures in L2 phonemic sequences: the case of Mandarin listeners' perceptual deletion of English post-vocalic laterals" *Phonetica*, advance online publication. https://doi.org/10.1515/phon-2022-0027.

Yang, Jing, and Robert A. Fox. 2017. "L1–L2 interactions of vowel systems in young bilingual Mandarin-English children." *Journal of Phonetics* 65: 60–76.

Yusa, N., K. Nasukawa, M. Koizumi, J. Kim, N. Kimura, and K. Emura. 2010. "Unexpected Effects of the Second Language on the First." New Sounds 2010: Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech: 580–584.

České resumé

Práce se zabývá produkcí a percepcí anglických vokálů /i, ɪ, ɛ, æ, ʌ/ českými dětmi ve věku 3-5 let, které jednou týdně navštěvují kurz angličtiny. Jejich výslovnost daných vokálů ve slovech byla nahrána šestkrát v intervalu tří měsíců. Cílem práce je popsat vliv dvou odlišných zdrojů inputu na produkci vokálů u dětí, které nemají přístup k interaktivnímu inputu od rodilých mluvčí angličtiny. Zejména u začátečníků může input s cizineckým přízvukem, který nezachovává kontrasty L2 vést k vývoji vokálů, které mohou být zdrojem nedorozumění při komunikaci v L2. Zkoumání realizace vokálů dětmi, které se učí cizí jazyk od nerodilých mluvčích, umožňuje popsat vliv inputu s cizineckým přízvukem na výslovnost dětí, a charakterizovat input, který by podpořil vývoj vokálů v L2 takovým způsobem, aby byla minimalizována možná nedorozumění. Práce se dále zabývala i percepcí vokálů /i, ɪ, ɛ, æ/ a fonetickým posunem v produkci českých vokálů.

Během hodin angličtiny byly děti vystaveny řeči mluvčí, která zachovávala kvalitativní kontrast mezi /æ, ɛ/. Mimo hodiny byly vystaveny anglickému inputu od českých mluvčích s cizineckým přízvukem, u kterých bylo předpokládáno, že kvalitativní kontrast mezi /ɛ, æ/ nedodržují. Byl formulován předpoklad, že děti si budou vokály z L2 spojovat s jinými vokály z L1 na základě realizace /ɛ, æ/ daným zdrojem inputu. Asimilace vokálů /i, ɪ, ʌ/ k vokálům L1 by měla podle zůstat stejná pro oba zdroje inputu. Na základě předpokladu, že děti si ukládají v mysli exempláře slov a s nimi spojené fonetické detaily jednotlivých realizací vokálů bylo očekáváno, že vliv zdroje inputu bude zřejmý na slovech, které děti slyšely primárně od jednoho ze dvou zkoumaných zdrojů inputu.

Výsledky potvrdily vliv zdroje inputu na produkci vokálu /æ/, který byl dvěma zdroji realizován odlišně. Ve slovech, které děti slýchali primárně od mluvčích mimo hodiny angličtiny, byly vokály /ɛ, æ/ realizovány jako velmi podobné jak ve výšce tak v retrakci (v ERB); ve slovech, které děti slýchaly v hodinách byly tyto vokály rozlišeny lépe. Na produkci /i, ɪ, ʌ/ neměl zdroj inputu významný vliv. Tyto výsledky zdůrazňují vliv zdroje inputu na realizaci anglických vokálů dětmi, které se učí angličtinu od nerodilých mluvčích. Odlišná realizace vokálů těmito mluvčími ovlivňuje realizaci těchto vokálů u dětí, a může tak přispět k rozvoji, nebo k potlačení formace nových kategorií pro vokály v L2.

Dále byl zkoumán vliv času na produkci vokálů v L2, realizace vokálů v prvních šesti týdnech byla porovnána s realizací vokálů v posledních šesti týdnech experimentu. Realizace většiny vokálů v L2 se výrazně nezměnila, pouze /i/ bylo vyslovováno jako vyšší v posledních šesti týdnech. Tato změna ale nepřispěla ke kvalitativnímu odlišení /ɪ/ od /i/, protože realizace vokálů se i nadále výrazně překrývala.

Výsledky percepčního experimentu odhalily vysokou individuální variabilitu při identifikaci chybně vyslovených slov v L2 s nahrazeným vokálem. Toto zjištění zvýrazňuje existenci mnoha faktorů, které mohou ovlivňovat výsledky předškolních dětí v percepčních experimentech.

Práce se také zabývala otázkou, jestli se vokály dětí v L1 mění, zatímco se učí L2; zaměřila se na produkci moravských vokálů /i:, i, ɛ, a/ ve slovech, které děti produkovaly ve dvou sezeních oddělených 10 týdny. Realizace vokálů se změnila, což naznačuje, že fonetický posun se vyskytuje i u nezkušených dětí, které se učí v prostředí, kde dominuje jejich mateřský jazyk.

Appendix

7.1 Characteristics of Czech and English monolinguals

Table 9. Characteristics of the Czech (left) and English (right) monolinguals: subject number, sex, age, (residence, source).

Subj. no.	Sex	Age	Subj. no.	Sex	Age	Residence	Source
11	F	4;1					
12	F	4;4	22	M	4;6	Bedfordshire	Gladwin (2020)[1]
13	F	4;4	23	F	3;9	Northamptonshire	Clark (2019a)[2]
14	F	5;3	24	M	5;0	Northamptonshire	Clark (2019b)[3]
15	F	5;4	25	F	5;5	Hertfordshire	Roskillys (2020)[4]
16	F	5;4	26	F	5;0	Greater London	Ashlee (2021)[5]
17	M	5;9	27	F	5;6	Berkshire	Mama (2021)[6]
18	F	5;11	28	F	5		Mrs (2017)[7]
19	M	6;3	29	M	5;1	Bedfordshire	Gladwin (2018)[8]
20	M	6;4	30	M	6	Bedfordshire	Gladwin (2019)[9]

7.2 Wordlists

Table 10. Lists of words recorded in English (left) and Czech (right). New words are bold.

Vowel	Word	IPA	Vowel	Word	IPA	Gloss
i	sheep	/ʃip/	i:	písek	/pi:sɛk/	sand
	beach	/bit͡ʃ/		sýr	/si:r/	cheese
	beet	/bit/		býk	/bi:k/	bull
	seed	/sid/		šíp	/ʃi:p/	arrow
	seat	/sit/				
ɪ	fish	/fɪʃ/	ɪ	tygr	/tɪgr/	tiger
	tick	/tɪk/		židle	/ʒɪdlɛ/	chair
	chips	/t͡ʃɪps/		kytka	/kɪtka/	flower

[1] Gladwin, Alex. 2020. "AN INTERVIEW WITH 4 YEAR OLD LITTLE BOY | Alex Gladwin." Uploaded January 15th 2020. Video, 6:46. https://www.youtube.com/watch?v=4xA5F_0s9jA.

[2] Clark, Kayleigh. 2019a. "INTERVIEW WITH MY 3 YEAR OLD | Kayleigh Clark." Uploaded May 2nd 2019. Video, 9:02. https://www.youtube.com/watch?v=07_-5zKI8xk.

[3] Clark, Kayleigh. 2019b. "INTERVIEW WITH MY 5 YEAR OLD | Kayleigh Clark." Uploaded January 3rd 2019. Video, 9:05. https://www.youtube.com/watch?v=Rw2W2R0K4jM.

[4] Roskillys, The. 2020. "Interview with my 5 year old." Uploaded October 3rd 2020. Video, 3:42. https://www.youtube.com/watch?v=kmNOEJzY9KE .

[5] Ashlee, Rhiannon. 2021. "Interview With My 5 Year Old & Cosy Girls Night | VLOGTOBER." Uploaded October 21st 2021. Video, 27:23. https://www.youtube.com/watch?v=DcAyfxRweUI.

[6] Mamma Striving High. 2021. "An Interview With My FIVE Year Old." Uploaded April 16th 2021. Video, 4:27. https://www.youtube.com/watch?v=eVvdaohFkQ0.

[7] Mrs Sausagefingers. 2017. "Interview with a 5 year old - Kids." Uploaded January 19th 2017. Video, 4:27. https://www.youtube.com/watch?v=QZpR-RRiBUE.

[8] Gladwin, Alex. 2018. "INTERVIEW WITH MY 5 YEAR OLD | Alex Gladwin." Uploaded May 23rd 2018. Video, 9:20. https://www.youtube.com/watch?v=CS4lICRPMdk.

[9] Gladwin, Alex. 2019. "AN INTERVIEW WITH MY 6 YEAR OLD | Alex Gladwin." Uploaded September 20th 2019. Video, 5:53. https://www.youtube.com/watch?v=jrmK4pPJgIE.

	dish	/dɪʃ/		šipka	/ʃɪpka/	pointer
	fist	/fɪst/	ɛ	pes	/pɛs/	dog
ɛ	bed	/bɛd/		šest	/ʃɛst/	six
	head	/hɛd/		led	/lɛt/	ice
	leg	/lɛg/		lev	/lɛf/	lion
	red	/ɹɛd/	a	hrad	/hrat/	castle
	text	/tɛkst/		krab	/krap/	crab
	dress	/dɹɛs/		had	/hat/	snake
	vet	/vɛt/		lak	/lak/	(nail) polish
	chest	/t͡ʃɛst/				
	vest	/vɛst/				
	chef	/ʃɛf/				
	step	/stɛp/				
	bread	/bɹɛd/				
æ	cat	/kæt/				
	back	/bæk/				
	hat	/hæt/				
	bat	/bæt/				
	sad	/sæd/				
	black	/blæk/				
	cap	/kæp/				
	happy	/hæpi/				
	stag	/stæg/				
	cash	/kæʃ/				
	rat	/ræt/				
	crab	/kɹæb/				
ʌ	bus	/bʌs/				
	hug	/hʌg/				
	dove	/dʌv/				
	hut	/hʌt/				
	cut	/kʌt/				

Table 11. Plan of recording sessions, teaching sessions and assessment

Week, Session	Data recorded after lesson	Words revised during lesson	Words taught
W0	Czech vocabulary and pronunciation assessment	Back, cat, hat, bed, leg, head, sheep, fish, bus, horse, dog, cars, foot, book, bird, shoes, grass, fast	-
W1	-	Back, cat, hat, bed, leg, head, sheep, fish, bus, horse, dog, cars, foot, book, bird, shoes, grass, fast	-
W2: S1	9 targets (cat, back, hat, leg, bed, head, sheep, fish, bus), other vowels (horse, dog, cars, foot, book, bird, skirt, shoes, grass, stars, fast)	Sad, red Pot, doors, goose, skirt, bird	-

		back, cat, hat, bed, leg, head, sheep, fish, bus, horse, dog, cats, foot, book, bird, skirt, shoes, grass, stars, fast	
W3	-	Sad, red Pot, doors, goose, skirt, bird	-
W4: S2	9 targets, sad, red Pot, doors, goose, bird, skirt Text, bat, seat, tick, hug	Sad, red Pot, doors, goose, skirt, bird	Text, bat, seat, tick, hug
W5	Czech production 1	Black, dress Text, bat, sea, tick, hug	-
W6	No class	No class	No class
W7	-	Black, dress Text, bat, seat, tick, hug	-
W8: S3	9 targets, black, dress Text, bat, seat, tick, hug Cap, vet, beach, chips, dove	Text, bat, seat, tick, hug Black, dress	Cap, vet, beach, chips, dove
W9	Perception session 1	Cap, vet, beach, chips, dove Chest, happy Perception words	-
W10		Cap, vet, beach, chips, dove Chest, happy	
W11: S4	9 targets, chest, happy Cap, vet, beach, chips, dove Stag, vest, beet, dish, hut	chest, happy Cap, vet, beach, chips, dove	Stag, vest, beet, dish, hut
W12	Perception session 2	Stag, vest, beet, dish, hut Rat, step	-
W13	-	Stag, vest, beet, dish, hut Rat, step	-
W14: S5	9 targets, rat, step Stag, vest, beet, dish, hut Cash, chef, seed, fish, cut	Stag, vest, beet, dish, hut Rat, step	Cash, chef, seed, fish, cut
W15	Czech production 2	Cash, chef, seed, fist, cut	

W16: S6	9 targets, crab, bread, cash, chef, seed, fist, cut	-	-

7.3 R analysis script

```
library(dplyr)
library(tidyr)
library(ggplot2)
library(tidyverse)
library(broom)
library(lme4)
library(readxl)
library(ggeffects)
library(emmeans)
library(RColorBrewer)
library(MuMIn)
library(openxlsx)
library(lmerTest)
library(phonR)

### load data

setwd("C:/Users/monika/Desktop/Kucerova_BT")
ppts <- read_excel("Kucerova_BT_data.xlsx")

# prepare data for analysis:
  # convert language, time, status, vowel to factors,
  # change english "ɪ" to "En_ɪ" and Czech "i" to "Cz_ɪ"
  ppts <- mutate(ppts, language = as.factor(language),
             time = as.factor(time),
             status = as.factor(status),
             vowel = as.factor(ifelse(phone == "ic", "ɪ",
                                  ifelse(phone == "is",
                                       ifelse(language == "Cs","ɪ", "i"),
                                       as.character(vowel)))))

  # analysis for only EFL children (all 9)
  efl <- filter(ppts, group == "efl")

  # analysing only front Vs
  efl_bothLs <- filter(efl, phone %in% c("a", "ae","e", "ef","vt", "is", "ic", "ii", "i"))

  # only english for the 9 efls
  efl_En <- filter(efl_bothLs, language == "En")

  # treatment code time, status (reference level = new)
  contrasts(efl_En$status) <- contr.treatment(2)
  contrasts(efl_En$time) <- contr.treatment(2)
  # time 2 as reference level
  efl_En = efl_En %>% mutate(time = relevel(time, "T2"))
```

```
  # ae as intercept
  efl_En = efl_En %>% mutate(vowel = relevel(vowel, 5))


### HEIGHT MODEL for English vowels
height.mdl <- lmer(F1e-F0e ~ vowel * time * status +
                  (1 + time | speaker)  + (1|word),
                data = efl_En,
                control = lmerControl(optimizer = "bobyqa"))

  summary(height.mdl)
  r.squaredGLMM(height.mdl)
  write.csv(round(summary(height.mdl)$coefficients, 3), "EN_height.csv")

  # check model residuals: q-q norm, q-q line, histogram
  res <- residuals(height.mdl)
  qqnorm(res)
  qqline(res)
  hist(res)

  # fitted plot
  height.estim <- ggemmeans(height.mdl, type = "fixed",
                      terms = c("status", "vowel", "time"), ci.lvl = 0.95)
  height.estim

  hifit <- plot(height.estim) +
    ggtitle("Predicted values of vowel height")
  hifit
  ggsave("hifit.png", plot = hifit, width = 4, height = 3, units = "in",
       dpi = 600)

  # individual speakers' deviations from the grand intercept (the summary(height.mdl)
intercept)
  ranef(height.mdl)$speaker
  write.csv((ranef(height.mdl)$coefficients), "EN_height_ranef.csv")

### RETRACTION MODEL for English vowels

retr.mdl <- lmer(F2e-F0e ~ vowel * time * status +
                (1 + time | speaker)  + (1|word),
              data = efl_En,
              control = lmerControl(optimizer = "bobyqa"))
  summary(retr.mdl)
  write.csv(round(summary(retr.mdl)$coefficients, 3), "EN_retraction.csv")
  r.squaredGLMM(retr.mdl)

  # individual speakers' deviations from the grand intercept
  ranef(retr.mdl)$speaker
```

```
  #fitted plot
  retr.estim <- ggemmeans(retr.mdl, type = "fixed",
                        terms = c("status", "vowel", "time"), ci.lvl = 0.95)
  retr.estim
  retrfit <- plot(retr.estim) +
    ggtitle("Predicted values of vowel retraction")
  retrfit
  ggsave("retrfit.png", plot = retrfit, width = 4, height = 3, units = "in",
         dpi = 600)

# PLOTS

# low vowels
efl_En_low <- filter(efl_bothLs, phone %in% c("vt","ae","ef"))
efl_En_low_old <- filter(efl_En_low, status == "old", time == "T2")
efl_En_low_new <- filter(efl_En_low, status == "new", time == "T2")

par(mfrow = c(1, 2))

### plot low English vowels in time 2: old vs new words
with(efl_En_low_old, plotVowels(F1e-F0e, F2e-F0e, vowel, group = status,
                              plot.tokens = TRUE, plot.means = TRUE, pch.means = vowel,
                              cex.means = 2, var.col.by = vowel, var.sty.by = vowel,
                              ellipse.line = TRUE,  ellipse.fill = FALSE, fill.opacity = 0.1,
                              xlim = c(19, 9),
                              ylim = c(12, 2),
                              xlab = "F2-F0 (in ERB)", ylab = "F1-F0 (in ERB)",
                              pch.tokens = vowel,
                              cex.tokens = 1, alpha.tokens = 0.4, col = c("red2",
                                                                    "springgreen4", "royalblue3"),
                              pretty = TRUE, main = "Old words in time 2"))

with(efl_En_low_new, plotVowels(F1e-F0e, F2e-F0e, vowel, group = status,
                              plot.tokens = TRUE, plot.means = TRUE, pch.means = vowel,
                              cex.means = 2, var.col.by = vowel,
                              ellipse.line = TRUE, ellipse.fill = FALSE, fill.opacity = 0.1, var.sty.by
= vowel,
                              xlim = c(19, 9),
                              ylim = c(12, 2),
                              xlab = "F2-F0 (in ERB)", ylab = "F1-F0 (in ERB)",
                              pch.tokens = vowel,
                              cex.tokens = 1, alpha.tokens = 0.4,
                              col = c("red2", "springgreen4", "royalblue3"),
                              pretty = TRUE, main = "New words in time 2"))


# high vowels: time 1 vs time 2
efl_En_high <- filter(efl_bothLs, phone %in% c("ic", "i")))

efl_En_high <- mutate(efl_En_high, vowel = as.factor(ifelse(phone == "ic", "ɪ",
```

```
                                    ifelse(phone == "i", "i",
                                        as.character(vowel)))))

efl_En_high1 <- filter(efl_En_high, time == "T1")
efl_En_high2 <- filter(efl_En_high, time == "T2")

# plot time 1
with(efl_En_high1, plotVowels(F1e-F0e, F2e-F0e, vowel, group = time,
                  plot.tokens = TRUE, plot.means = TRUE, pch.means = vowel,
                  cex.means = 2, var.col.by = vowel, var.sty.by = vowel,
                  ellipse.line = TRUE, ellipse.fill = FALSE, fill.opacity = 0.1,
                  xlim = c(20, 13),
                  ylim = c(6, -1),
                  xlab = "F2-F0 (in ERB)", ylab = "F1-F0 (in ERB)",
                  pch.tokens = vowel,
                  cex.tokens = 1, alpha.tokens = 0.4,
                  pretty = TRUE, main = "Time 1"))

# plot time 2
with(efl_En_high2, plotVowels(F1e-F0e, F2e-F0e, vowel, group = time,
                  plot.tokens = TRUE, plot.means = TRUE, pch.means = vowel,
                  cex.means = 2, var.col.by = vowel, var.sty.by = vowel,
                  ellipse.line = TRUE, ellipse.fill = FALSE, fill.opacity = 0.1,
                  xlim = c(20, 13),
                  ylim = c(6, -1),
                  xlab = "F2-F0 (in ERB)", ylab = "F1-F0 (in ERB)",
                  pch.tokens = vowel,
                  cex.tokens = 1, alpha.tokens = 0.4,
                  pretty = TRUE, main = "Time 2"))

### CZECH VOWELS produced by the 7 learners: time 1 vs time 2

 # filter only Czech data for 7 learners, short front Vs and i:
 dataCz <- filter(ppts, language == "Cs", speaker %in%
c("s1","s4","s5","s6","s7","s8","s9"))
 dataCz <- filter(dataCz, phone %in% c("a", "e", "is", "ii"))

 dataCz <- mutate(dataCz, vowel = as.factor(ifelse(phone == "e", "ɛ",
                                    ifelse(phone == "ef", "ɛ",
                                        as.character(vowel)))))

 # /ɛ/ as intercept
 dataCz = dataCz %>% mutate(vowel = relevel(vowel, 2))

### HEIGHT of Czech vowels
heightCz.mdl <- lmer(F1e-F0e ~ vowel * time +
                (1+time|speaker)  + (1|word),
               data = dataCz,
               control = lmerControl(optimizer = "bobyqa"))
```

```
  summary(heightCz.mdl)
  write.csv(round(summary(heightCz.mdl)$coefficients, 3), "heightCz.csv")
  r.squaredGLMM(heightCz.mdl)

  # check model residuals: q-q norm, q-q line, histogram
  res <- residuals(heightCz1.mdl)
  qqnorm(res)
  qqline(res)
  hist(res)

  #fitted plot
  heightCz.estim <- ggemmeans(heightCz.mdl, type = "fixed",
                        terms = c("vowel", "time"), ci.lvl = 0.95)
  heightCz.estim
  hi_Cz_fit <- plot(heightCz.estim) +
    ggtitle("Predicted values of L1 vowel height at time 1 and time 2")
  hi_Cz_fit
  ggsave("hi_Cz_fit.png", plot = hi_Cz_fit, width = 6, height = 4, units = "in",
       dpi = 800)


### RETRACTION of Czech vowels
retrCz.mdl <- lmer(F2e-F0e ~ vowel * time +
                     (1+time|speaker)  + (1|word),
                   data = dataCz,
                   control = lmerControl(optimizer = "bobyqa"))

  summary(retrCz.mdl)
  write.csv(round(summary(retrCz.mdl)$coefficients, 3), "retrCz.csv")
  r.squaredGLMM(retrCz.mdl)

  # check model residuals: q-q norm, q-q line, histogram
  res <- residuals(heightCz1.mdl)
  qqnorm(res)
  qqline(res)
  hist(res)

  #fitted plot
  retrCz.estim <- ggemmeans(retrCz.mdl, type = "fixed",
                       terms = c("vowel", "time"), ci.lvl = 0.95)
  retrCz.estim
  retr_Cz_fit <- plot(retrCz.estim) +
    ggtitle("Predicted values of L1 vowel retraction at time 1 and time 2")
  retr_Cz_fit
  ggsave("retr_Cz_fit.png", plot = retr_Cz_fit, width = 6, height = 4, units = "in",
       dpi = 800)

# PLOT Czech vowels

  # only Czech short front Vs + i: from 7 learners
```

```
  cz_vowels <- filter(ppts, language == "Cs", group == "efl", phone %in% c("is", "ii",
"e","a"), speaker %in% c("s1","s4","s5","s6","s7","s8","s9"))
  czech <- filter(ppts, group == "czech", phone %in% c("is","ii", "e","a"))
  cz_vowels <- rbind(czech, cz_vowels)

  # change symbols
  cz_vowels <- mutate(cz_vowels, vowel = as.factor(ifelse(phone == "is", "ɪ",
                                              ifelse(phone == "ii", "iː",
                                                     as.character(vowel)))))
  cz_vowels <- mutate(cz_vowels, vowel = as.factor(ifelse(phone == "e", "ɛ",
                                                ifelse(phone == "ef", "ɛ",
                                                       as.character(vowel)))))
 # time 1 vs time 2
  cz_vowels1_efl <- filter(cz_vowels, group == "efl",  time == "T1")
  cz_vowels1_czech <- filter(cz_vowels, group == "czech")
  cz_vowels1 <- rbind(cz_vowels1_czech, cz_vowels1_efl)
  cz_vowels2_efl <- filter(cz_vowels, group == "efl",  time == "T2")
  cz_vowels2 <- rbind(cz_vowels1_czech, cz_vowels2_efl)

# plot time 1
  with(cz_vowels1, plotVowels(F1e-F0e, F2e-F0e, vowel, group= group,
                      plot.tokens = FALSE, plot.means = TRUE, pch.means = vowel,
                      cex.means = 2, var.col.by = group, var.sty.by = group,
                      ellipse.line = TRUE,  ellipse.fill = FALSE, fill.opacity = 0.1,
                      # xlim = c(22, 10),
                      #ylim = c(10, 0.5),
                      xlab = "F2-F0 (in ERB)", ylab = "F1-F0 (in ERB)",
                      pch.tokens = vowel,
                      cex.tokens = 1, alpha.tokens = 0.4,
                      #col = c("red2", "springgreen4", "royalblue3"),
                      legend.kwd = "bottomleft",
                      pretty = TRUE, main = "Czech vowels produced by learners (time 1) and
monolinguals"))

# plot time 2
  with(cz_vowels2, plotVowels(F1e-F0e, F2e-F0e, vowel, group= group,
                      plot.tokens = FALSE, plot.means = TRUE, pch.means = vowel,
                      cex.means = 2, var.col.by = group, var.sty.by = group,
                      ellipse.line = TRUE,  ellipse.fill = FALSE, fill.opacity = 0.1,
                      # xlim = c(22, 10),
                      #ylim = c(10, 0.5),
                      xlab = "F2-F0 (in ERB)", ylab = "F1-F0 (in ERB)",
                      pch.tokens = vowel,
                      cex.tokens = 1, alpha.tokens = 0.4,
                      #col = c("red2", "springgreen4", "royalblue3"),
                      legend.kwd = "bottomleft",
                      pretty = TRUE, main = "Czech vowels produced by learners (time 2)
and monolinguals"))

# plot English vowels produced by learners vs by SSBE monolinguals
```

```
# filter speakers and vowels
# use time 2 and status new for learners
  monos_en <-  filter(ppts, phone %in% c("i", "ic", "ef", "ae","vt"), group == "ssbe")
  efl_T2_new <- filter(ppts, phone %in% c("i", "ic", "ef", "ae","vt"), group == "efl", time
=="T2", status =="new")
  monos_efl <- rbind(monos_en, efl_T2_new)
  monos_efl <- mutate(monos_efl, vowel = as.factor(ifelse(phone == "e", "ɛ",
                                                 ifelse(phone == "ef", "ɛ",
                                                      as.character(vowel)))))

# plot learners vs SSBE monolinguals
  with(monos_efl, plotVowels(F1e-F0e, F2e-F0e, vowel, group= group,
                  plot.tokens = FALSE, plot.means = TRUE, pch.means = vowel,
                  cex.means = 2, var.col.by = group, var.sty.by = group,
                  ellipse.line = TRUE,  ellipse.fill = FALSE, fill.opacity = 0.1,
                  # xlim = c(22, 10),
                  #ylim = c(10, 0.5),
                  xlab = "F2-F0 (in ERB)", ylab = "F1-F0 (in ERB)",
                  pch.tokens = vowel,
                  cex.tokens = 1, alpha.tokens = 0.4,
                  #col = c("red2", "springgreen4", "royalblue3"),
                  legend.kwd = "bottomleft",
                  pretty = TRUE, main = "English Vs produced by learners (new, time 2) and
monolinguals"))


# plot teacher and adult monolinguals
# load teacher's data
  ex <- read_excel("reference_teacher.xlsx")

# change symbols
  ex <- mutate(ex, vowel = as.factor(ifelse(phone == "e", "ɛ",
                                ifelse(phone == "i",
                                     ifelse(language == "Cs","ɪ", "i"),
                                     as.character(vowel)))))

# filter out back vowels
  ex <- filter(ex, phone %in% c("ic", "i", "vt", "ae", "ef", "ii", "a","e"))

  par(mfrow = c(1,2))

# plot teacher's Czech and English productions
  with(ex, plotVowels(F1e-F0e, F2e-F0e, vowel, group = language,
                plot.tokens = FALSE, plot.means = TRUE, pch.means = vowel,
                cex.means = 2, var.col.by = language,
                ellipse.line = FALSE, ellipse.fill = FALSE, fill.opacity = 0.1,
                xlim = c(18, 12),
                ylim = c(9, 2),
                xlab = "F2-F0 (in ERB)", ylab = "F1-F0 (in ERB)",
                pch.tokens = vowel,
```

```
                  cex.tokens = 1, alpha.tokens = 0.4, col = c( "red2", "black"),
                  legend.kwd = "bottomleft",
                  pretty = TRUE, main = "English and Czech vowels produced by the
teacher"))

  # load data on adult monolinguals
  ref <- read_excel("reference_adult_monolinguals.xlsx")
  ref <- filter(ref, source %in% c("SkaVol12", "Bje16"))

# plot adult SSBE vs Czech monolinguals
  with(ref, plotVowels(F1, F2, vowel, group = source, log = "xy",
                  plot.tokens = FALSE, plot.means = TRUE, pch.means = vowel,
                  cex.means = 2, var.col.by = source,
                  ellipse.line = FALSE, ellipse.fill = FALSE, fill.opacity = 0.1,
                  xlab = "F2 (in Hz)", ylab = "F1 (in Hz)",
                  pch.tokens = vowel,
                  ylim = c(800, 250),
                  cex.tokens = 1, alpha.tokens = 0.4,
                  legend.kwd = "bottomleft",
                  pretty = TRUE, main = "English and Czech vowel frequencies from
monolingual speakers", col = c( "black", "red2"),))

### PERCEPTION
  # load data
perc_updated <- read.xlsx("perception_data.xlsx")

# separate learners into two groups
par(mfrow = c(1,2))
perc_g <- filter(perc_updated, speaker %in% c("AV", "CT", "ES", "KA", "SS"))

perc.count2 <- as_tibble(with(perc_g, table(correctV, reaction, congruence)))

# plot group 1
perc_g.plot <- ggplot(perc.count2, aes(x=correctV, y=n, fill=reaction))+
  ggtitle("Group 1: percentage of 'accept' reactions") +
  facet_grid(.~factor(congruence, levels = c("originalV", "replacedV")))+
  geom_bar(position="fill",stat = "identity") +
  theme_light() + ylab("")+
  scale_y_continuous(breaks = seq(0,1,.25),
                 labels = scales::percent)+
  scale_fill_manual("Reaction",
                 labels=c( "accept", "reject"),
                 values = c("#306100", "brown1"))
perc_g.plot

# plot group 2
perc_b <- filter(perc_updated, speaker %in% c("AI", "JA", "MN", "VD"))

perc.count_b <- as_tibble(with(perc_b, table(correctV, reaction, congruence)))
```

```
perc_b.plot <- ggplot(perc.count_b, aes(x=correctV, y=n, fill=reaction))+
  ggtitle("Group 2: percentage of 'accept' reactions")+
  facet_grid(.~factor(congruence, levels = c("originalV", "replacedV")))+
  geom_bar(position="fill",stat = "identity") +
  theme_light() + ylab("")+
  scale_y_continuous(breaks = seq(0,1,.25),
                     labels = scales::percent)+
  scale_fill_manual("Reaction",
                    labels=c( "accept", "reject"),
                    values = c("#306100", "brown1"))
perc_b.plot
```

7.4 Language background questionnaire

This questionnaire regarding language background of English learners created on the basis of the Czech version of LEAP-Q (Blumenfeld and Kaushanskaya 2007a) and the original English version of LEAP-Q for parents (Blumenfeld and Kaushanskaya 2007b). Questions and answers were translated into English by the author of this thesis.

(1) Please list all of the languages your child knows in order of dominance.

(2) Please list all of the languages your child knows in order of acquisition (their native language first).

(3) Please list what percentage of the time your child currently on average (passively) hears each language: (Percentages should add up to 100%).

(4) Please give the age (in years and moths) at which your child was enrolled in these English lessons.

(5) Please list any other language lessons your child attends at present: specify the form of the lessons, their frequency and duration of one lesson.

(6) Please list any other language lessons your child attended in the past: specify the form of the lessons, the age at which your child started and stopped attending the lessons, their frequency and duration of one lesson.

(7) Please give the approximate frequency and duration of child's exposure to the audio recordings provided with these language lessons per week.

(8) How frequently and how long does your child actively interact in English with a member of his/her family?

(9) How frequently and how long does your child interact in English with friends?

(10) Does your child like interacting in English? (Yes/No)

(11) Please state the duration your child has spent in an English-speaking environment (years and months): English speaking country, English-speaking family, English-speaking kindergarten

(12) Please circle how much the following factors contributed to your child learning English (options range from 0 = “did not contribute at all” to 10 = “contributed the most”).

(12) a. Interaction with friends

(12) b. Interaction with family

(12) c. English audio recordings

(12) d. Watching TV

(12) e. Watching YouTube and/or videos on the internet

(12) f. Playing computer/mobile games

(12) g. Any other influences you wish to mention

(13) How strong do you consider the foreign accent of the person who speaks English with your child the most (on a scale of 0 to 10, 0 = “none”, 10 = “pervasive”)

(14) Does your child have a vision problem __, hearing impairment __, language disability __, or learning disability __? (Check all applicable). If yes, please explain (including any corrections).

Additional questions regarding dialect

(15) Is your child currently exposed to any dialect?

(16) If you answered yes to the previous question, please specify the proportion of time (0-100%) your child is exposed to the dialect, and give the name of the dialect or specify a geographical location which it is typical for.

(17) Does you child speak any dialect?

(18) If you answered yes to the previous question, which dialect from question (16) does your child use and specify the proportion of time he/she uses the dialect (0-100%) relative to common Czech.

(19) If your child speaks a dialect, in which environment does he/she speak it?

(19) a. In kindergarten

(19) b. At home

(19) c. With friends outside kindergarten

(19) d. When talking to older people (e.g. grandparents)

(19) e. In other circumstances, specify.

Q	s7	s9	s2	s1	s6	s3	s8	s10	s4	s5
1	Czech, English	Czech, English	Czech, English, Russian (Belarus)	Czech, English	Czech, English	Czech, English	Czech, Slovak, English	Czech, English	Czech, English	Czech, English
2	Czech, English	Czech, English	Czech, English, Russian (Belarus)	Czech, English	Czech, English	Czech, English	Slovak, Czech, English	Czech, English	Czech, English	Czech, English
3	Czech 90%, English 10%	Czech 95%, English 5%	Czech 60%,English 35%, Russian 5%	Czech 90%, English 10%	Czech 90%, English 10%	Czech 98%, English 2%	Czech 70%, Slovak 30%, English 5%	Czech 95%, English 5%	Czech 95%, English 5%	Czech 90%, English 10%
4	3 years old	3 years, 3 moths	4 years, 2 moths	3 years, 3 months	1 year, 9 months	2 years, 6 months	3 years old	3 years, 2 months	2 years, 10 months	2 years, 10 months
5	-	-	-	-	-	-	-	-	English lessons in kindergarten, 45 minutes a week	-
6	-	-	-	-	-	-	-	-	-	-
7	10 minutes 3 times a week	1 hour weekly	20 minutes every day	10 minutes every day	15 minutes twice a week	15 minutes every day	10 minutes three times a week	Is not exposed to audio recordings	5 minutes once a week	30 minutes once a week
8	30 minutes 3 times a week	-	2 hours a week	20 minutes twice a week	10 minutes two times a week	Occasionally in conversation	10-15 minutes once or twice a week	10 minutes a week	-	-
9	-	-	-	-	-	-	-	-	-	-
10	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
11	-	-	-	-	-	-	-	-	-	-
12 a.	5	1	0	0	0	1	0	0	3	0

12 b.	5	1	7	7	3	4	3	3	0	0
12 c.	5	5	7	7	4	7	4	0	8	5
12 d.	5	0	7	6	4	7	9	4	8	1
12 e.	5	4	0	6	4	7	9	1	8	5
12 f.	5	0	0	8	0	0	2	0	8	1
12g.	-	8 phone apps	English teacher	-	-	-	-	-	-	-
13	5	5	5	4	5	6	7	10	-	0
14	-	-	-	-	-	-	-	-	-	-
15	Yes	No	No	No	No	No	No	No	No	No
16	70% hanáčtina	-	-	-	-	-	-	-	-	-
3	No	No	No	No	No	No	No	No	No	No
4	-	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-

Anotace

Autor: Monika Kučerová
Studijní obor: Anglická filologie
Název: Vývoj anglických vokálů u českých dětí navštěvujících kurzy angličtiny
Typ práce: bakalářská práce
Fakulta a katedra: Filozofická fakulta, Katedra anglistiky a amerikanistiky
Vedoucí práce: Mgr. Šárka Šimáčková, Ph.D.
Počet stran: 49
Počet znaků: 87710
Charakteristika: Práce se zabývá produkcí a percepcí anglických vokálů /i, ɪ, ɛ, æ, ʌ/ českými dětmi ve věku 3-5 let, které se učí anglicky. Výslovnost vokálů ve slovech byla nahrána šestkrát v intervalu tří měsíců. Cílem bylo popsat vliv dvou odlišných zdrojů inputu na produkci vokálů u dětí. Zejména u začátečníků může input s cizím přízvukem, který nezachovává kontrasty L2 vést k vývoji vokálů, které budou zdrojem nedorozumění při komunikaci v L2. Výsledky potvrdily vliv zdroje inputu na produkci vokálu /æ/. Ve slovech, které děti slýchali primárně od mluvčích mimo hodiny angličtiny byly vokály /æ, ɛ/ realizovány jako kvalitativně podobné; ve slovech, které děti slýchaly v hodinách byly tyto vokály lépe rozlišeny. Na produkci /i, ɪ, ʌ/ neměl zdroj inputu významný vliv. Dále byl zkoumán vliv času na produkci vokálů v L2, byla srovnána jejich realizace v prvních a posledních šesti týdnech experimentu. Pouze /i/ se změnil v čase, byl vyšší v posledních šesti týdnech. Tato změna ale nepřispěla k odlišení /i/ od /ɪ/, protože se i nadále výrazně překrývaly. Práce se dále zabývala i percepcí vokálů /i, ɪ, ɛ, æ/. Výsledky percepčního experimentu odhalily vysokou individuální variabilitu při identifikaci chybně vyslovených slov v L2 s nahrazeným vokálem. Dále se práce věnovala fonetickému posunu v moravských vokálech /i:, i, e, a/ ve slovech, které děti produkovaly ve dvou sezeních oddělených 10 týdny. Realizace vokálů se změnila, což naznačuje, že fonetický posun se vyskytuje i u nezkušených dětí, které se učí cizí jazyk v prostředí, kde dominuje jejich mateřský jazyk.

Klíčová slova: osvojování cizího jazyka, anglické vokály, fonetický posun, děti předškolního věku, cizí přízvuk