

Filozofická fakulta Univerzity Palackého

**Subconscious Imitation of Phonetic Features of
Perceived Speech and its Influence on
Phonological Contrast**

(Bachelor Thesis)

2023

Anna Kovaříková

Filozofická fakulta Univerzity Palackého

Katedra anglistiky a amerikanistiky

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

V Olomouci dne 9. 5. 2023

Anna Kovaříková

Acknowledgement

I would like to thank my supervisor Mgr. Václav Jonáš Podlipský, Ph.D. for his guidance, help and useful advice throughout the process of writing and for providing me with the materials and technology necessary to complete this thesis.

Abstract

Speakers spontaneously and subconsciously imitate the accents of their interlocutors. This thesis reviews and elaborates on prior research related to the relationship between the perception and production of phonetic properties of speech sounds and how it is influenced by the need to maintain phonological contrast between phonemes and other factors. Specifically, this thesis asks whether speakers imitate long Czech vowel /u:/ duration which was manipulated to either enhance or threaten the phonological contrast between the long vowel /u:/ and the short Czech vowel /u/ despite the fact that the stimulus word follows a short introductory phrase with a fixed speaking tempo. While prolonging the duration of the long vowel /u:/ enhanced the phonological contrast, the shortening threatened it. This thesis carried out a shadowing experiment where disyllabic stimuli words followed five-syllable carrier phrases. The results of the experiment suggest that speakers do strive to maintain phonological contrast as the participants imitated the prolonged /u:/ durations, but did not imitate the reduced /u:/ durations.

Table of Contents

Abstract	Chyba! Záložka není definována.
1 Introduction.....	7
2 Phonetic Imitation Background	8
2.1 Theories of Memory Systems.....	9
3 Factors Influencing Phonetic Imitation.....	10
3.1 Age.....	11
3.2 Gender	12
3.3 Social Selectivity and Opinion	13
3.4 Social Engagement	15
3.5 Conversational Dominance.....	16
3.6 Task Engagement	17
3.7 Exposure	18
3.8 Word Frequency	18
3.9 Phonological Neighbourhood Density.....	20
3.10 Number of Syllables in Words	20
3.11 Language Distance	21
3.12 Contrast Maintenance and Perceptual Salience.....	22
4 The Present Study	24
4.1 Methods	24
4.1.1 Material	25
4.1.2 Procedure.....	25
4.1.3 Participants	26
4.2 Results	27
5 Discussion	30
6 Conclusion	33
7 Appendices.....	34
7.1 Appendix 1: List of target words with their English translations.....	34
7.2 Appendix 2: List of carrier phrases with their English translations	35
8 Works Cited	36

9	Annotation.....	40
10	Anotace v češtině	41

1 Introduction

The phenomenon of phonetic imitation in speech has been analysed in a number of studies in the field of phonetics and phonology and it continues being investigated even today as many questions remain unanswered. One of the studied topics is concerned with the factors which influence phonetic imitation. So far, many factors have been proposed to influence the probability and degree of imitation. It has been suggested that children tend to imitate their interlocutors more than adults (Nielsen, 2014) and that speakers with the same first-language dialect tend to imitate each other more than speakers with different dialects (Kim et al., 2011). Researchers found that a positive opinion of a speaker (Babel et al., 2013) and the perceived attractiveness of a speaker (Babel et al., 2014) also encourage the imitation of the interlocutors' phonetic characteristics. Moreover, there are also indications that the factors of perceptual salience (Podlipský and Šimáčková, 2015) or contrast preservation (Nielsen, 2011) can be at play. These are just a selected few.

Besides phonetic imitation, this thesis is concerned with the factor of contrast preservation. To be more specific, I would like to answer the question of whether native Czech speakers will imitate differences in vowel duration to the point of phonological contrast between long and short vowels being threatened. The basis of this paper will be set in the study of Podlipský and Šimáčková (2015) who tested whether native Czech speakers are equally likely to imitate the reduction and the extension of vowel duration and of a stop prevoicing even though the reduction threatens and the extension enhances the preservation of phonological contrasts. Their hypothesis that the speakers would try to preserve the phonological contrast and avoid reduction held for prevoicing in stops but it did not hold for the duration of long vowels. As the reduction of the duration of long Czech vowels is perceptually salient to native listeners, Podlipský and Šimáčková (2015) concluded that contrast preservation does not completely prevent imitation.

However, Kim and Clayards (2019) later suggested that the imitation of both extended and reduced vowel duration in Podlipský and Šimáčková's research (2015) could be caused by the fact that 'listeners interpreted [the stimuli] not as shortened and extended vowels but as vowels spoken with a fast or slow speaking rate' (Kim & Clayards, 2019, p. 783). This thesis aims to replicate Podlipský and Šimáčková's experiment although with a slightly different method to rule out the possibility of mistakes in interpretation by the listeners, as suggested by Kim and Clayards (2019).

The theoretical part of this thesis focuses on phonetic imitation. The topic of phonetic imitation is introduced using the Communication Accommodation Theory by Giles et al. (1991). After that, I introduce two basic views on storing detailed phonetic information of words and a description of the proposed factors influencing phonetic imitation. Following the theoretical part is a detailed description of the experiment which was carried out as part of the thesis and its results. I describe the methodology of the experiment and the reasons for using such methods in the current study. The experiment carried out consisted of five phases including baseline recording, two shadowing tasks in which the participants repeated words from manipulated recordings and two post-shadowing recordings similar to the baseline recording. The methodology is based on Podlipský and Šimáčková's experiment (2015) which was modified to include the context of short introductory phrases to rule out the possibility of perception differences in speech tempo by the listener as was suggested by Kim and Clayards (2019). At the end, I sum up the results of the experiment and I draw conclusions based on the findings of the experiment.

2 Phonetic Imitation Background

This chapter focuses on introducing the topic of phonetic imitation using some of the most prominent theories of memory regarding phonetic detail, after which I introduce the factors which influence phonetic imitation and relevant research concerned with them.

Phonetic imitation is the phenomenon where the speaker's pronunciation becomes more similar to the recently perceived pronunciation of another speaker. It is also sometimes referred to as phonetic convergence, speech accommodation, phonetic alignment, or entrainment (Pardo, 2017). Phonetic imitation happens due to exposure. When a speaker is in a dialogue with another speaker, they may subconsciously imitate each other.

A prominent theory regarding imitation in communication is the Communication Accommodation Theory. The first signs of this theory emerged in a paper by Giles (1973). In his paper, Giles introduced a model of phonetic imitation which involves two types of a phenomenon which he called accent mobility: accent divergence, and accent convergence. This model became the main point of the theory dubbed 'the speech accommodation theory'. Giles et al. (1987) proposed that the theory change its name to 'communication accommodation theory' due to the theory expanding beyond just language.

Communicative Accommodation Theory seeks to theoretically describe the accommodation of communication, its motives and its consequences (Giles et al., 1991). CAT

describes imitation as a result of speakers' needs for 'social integration or identification with another' (Giles et al., 1991, p. 18). This hypothesis was derived from the similarity-attraction theory by Byrne (1971) and since then, it became one of the most accepted reasons for phonetic imitation. Byrne's theory proposes that attraction highly correlates with similarity in opinions and values. For example, in Babel (2012), female participants were found to imitate vowels of the speaker who was previously rated as attractive. Imitation could therefore be a means of becoming more similar and gaining social approval which would cause the speaker to be more likeable (Giles et al., 1991). Another motive for imitation proposed by CAT is 'a desire to regulate comprehension and increase communicative efficiency' (Thakerar et al., 1982, as cited in Dragojevic et al., 2016). Imitating another's speech would therefore make communication easier for the speakers. As Dragojevic et al. (2016) mention, motivation is established as the main force driving communication accommodation, no matter whether it is conscious or not. Besides imitation, divergence is also explained in terms of CAT as a means to emphasise differences between speakers (Giles et al., 1991). Divergence was demonstrated in Bourhis and Giles (1977), where Welsh speakers diverged more when they spoke to a person who expressed an opinion which threatened the Welsh identity.

In their work, Giles et al. describe imitation as a process during which 'individuals adapt to each other's communicative behaviours in terms of a wide range of linguistic-prosodic-nonverbal features' (1991, p. 7). In the quote, Giles et al. describe imitation as a spontaneous phenomenon. It is, however, necessary to mention the discrepancy in terms when it comes to phonetic imitation. Some research papers use the term 'phonetic imitation' while describing conscious imitation. In comparison, the term 'convergence' is mostly used while describing subconscious imitation. Sometimes the terms which describe subconscious or conscious imitation are used interchangeably in various research papers and there is not a clear consensus on the use or a clearly defined line between terms. This thesis focuses on subconscious imitation, although some experiments mentioned are concerned with instructed imitation. Throughout the thesis, I will mostly use the term 'imitation' and, unless explicitly stated, this term will stand for subconscious imitation.

2.1 Theories of Memory Systems

In order to study phonetic imitation in detail, knowledge of memory structure and signal processing is necessary as it could influence the potential factors for imitation. There has long been a debate on how words are stored in speakers' minds and depending on which theory

would be correct, the outcome would shape our view of the process of phonetic imitation and the mental processes which lead to imitation.

The two basic views emerging from this debate are often called abstractionist and exemplar views (Goldinger, 1996; Goldinger, 1998). These theories debate whether or not phonetic detail is stored in long-term representations of phonological vocalic or consonantal categories. Both of these views have fuelled a significant amount of research in relation to phonetic imitation.

The abstractionist view proposes that phonetic detail is not stored in episodic memory (Posner, 1964). Abstractionist theories claim that once incoming acoustic signals are processed and the speech input is parsed for linguistic content, the phonetic detail is discarded and does not enter abstract phonological representations forming part of entries in the mental lexicon.

Opposing the abstractionist view is the exemplar, also called the episodic, view. Theories within the exemplar view claim that phonetic detail is available within episodic memory functions within the lexical representations of individual words. They hypothesise that words in the mental lexicon are represented by accumulated memories of specific acoustic signals (also known as exemplars) not stripped of their phonetic detail (Goldinger, 1996). Perception as described by exemplar theories involves a signal in the stored collection being activated by a perceived acoustic signal. If the perceived word is familiar or more frequent, the number of activated signals is higher. This is connected to word frequency being considered an influence on phonetic imitation (see 3.8). Words that are less frequent have fewer stored signals which can be activated; therefore, the newly perceived token has potentially more influence and can lead to greater possibility or degree of imitation. The exemplar theory hypothesises that the activated signals all contribute to producing a word (Goldinger, 1996). As Babel (2012) explains, this means that phonetic imitation is seen as cumulative, the more the participants are exposed to a phenomenon, the more likely they are to imitate it (see 3.7, for the factor of Exposure).

3 Factors Influencing Phonetic Imitation

This chapter focuses on the factors which were studied in relation to potential influences on phonetic imitation. The following sections review studies which have been mainly or partially concerned with factors that influence phonetic imitation.

3.1 Age

The first factor I will focus on is the age of the speaker. In 2014, Nielsen found that children were more likely to imitate voice onset time (VOT) values in stop consonants than adults. Her study consisted of three age groups of speakers of American English which were examined: preschoolers, third graders and adult college students. The experiment consisted of a baseline recording, an exposure phase where the participants listened to the stimuli and a post-test which examined the influence of the stimuli. While all groups imitated the prolonged VOT (i.e., longer aspiration of voiceless stops), the two younger groups imitated greater than the adults. Nielsen (2014) hypothesised that the results are likely connected to the phonological representations of children not yet being fully developed. Nielsen concluded that the results are supportive of exemplar-based theories as they would argue that children would have fewer exemplars due to them not being exposed to as much speech as adults, therefore the stimuli would be more influential.

Nielsen's results (2014) contrast with the results of Paquette-Smith et al. (2022) who found no difference between children and adults in their experiment. Paquette-Smith et al. modelled their experiment after Nielsen (2014) with speakers of Canadian English. However, no imitation was found in their recordings. Paquette-Smith et al. proposed that perhaps the reason for the contrasting results was due to their baseline recordings of the child participants having longer VOT when compared to the baseline recordings by Nielsen (2014). They concluded that the children in their experiment likely hyperarticulated words or had longer VOTs overall.

While Nielsen (2014) and Paquette-Smith et al. (2022) focused on studying phonetic imitation of young children and adults, Schertz and Johnson (2022) tried comparing VOT imitation by adult and teenage speakers of Canadian English who were explicitly instructed to imitate. The authors pointed out that while many papers on adults and children and their imitation came out, the area of teenage imitation was rather unexplored. Their experiment consisted of an exposure phase in which the participants were listening to both manipulated and original VOT and the imitation phase where they were instructed to imitate. They found that there is not much difference in the imitation of VOT between adults and teenagers, the adults' imitation shift was found to be only slightly larger than that of the teenagers. Schertz and Johnson suggested two possibilities explaining why adults imitated more. They proposed that adults imitate abstract targets more overall or that adults are 'more likely to consider abstract properties of the stimulus as the target of imitation' (Schertz and Johnson, 2022, p. 1846).

3.2 Gender

The factor of gender was studied in relation to both sides of imitation: the imitating individual and the imitated individual. Namy et al. (2002) conducted an experiment concerned with the influence of gender on phonetic imitation. The participants took part in a shadowing task. Namy et al. measured no specific phonetic property, but the overall utterance imitation. They based the results on a perceived imitation reported by an AXB perceptual test where a separate group of listeners were presented with the shadowed utterance, the baseline utterance and the utterance of the model speaker, and were asked to choose the most similar-sounding utterance to the shadowed utterance. The gender of shadowing individuals and the gender of model speakers were studied. The results showed that, overall, female shadowing participants imitated more than male shadowing participants and that participants imitated male speakers more than female speakers. Namy et al. suggested that these results may be caused by the fact that ‘women are differentially socialized or reinforced throughout development to attend to indexical features such as emotional tone of voice’ (2002, p. 428).

Pardo’s research results (2006) contradicted the findings of Namy et al. (2002). Pardo compared how the role and gender of the talker in a pair affected imitation. She conducted a task-based experiment. Participants were sorted into pairs and each of them was assigned a map with various landmarks. One of the maps assigned to the pair had an additional path drawn on it which the participants were tasked to duplicate on the second map by communicating with each other. The results were analysed in a perceptual similarity test where a separate set of listeners judged the imitation in the produced utterances. No specific phonetic property was considered. Pardo found that male pairs imitated more than female pairs during the experiment. Pardo’s results (2006) were later replicated by Pardo et al. (2010). However, in the study by Pardo et al. (2010), the participants were explicitly instructed to imitate. The results of Pardo et al. (2010) suggested that males imitated more than females overall. Pardo et al. (2010) also addressed the hypothesis of Namy et al. (2002) in which Namy et al. suggested that women learn to pay more attention to indexical features which would make them more likely to imitate. Pardo et al. (2010), in contradiction, hypothesised that attention might be the decisive factor rather than women’s sensitivity to indexical features and that men and women might have different predispositions which influence when they are more likely to pay attention.

Nielsen (2014) also concluded her research concerned with VOT imitation with a conclusion that the factor of gender did not significantly influence the degree of phonetic imitation. She conducted a modified picture-naming experiment with speakers of American

English differentiated into three age groups: pre-schoolers, third graders, and college students. The participants took part in three sessions: a baseline session where they were shown pictures and they were asked to name what was on the picture, a shadowing session and a post-test which consisted of the same procedure as the baseline session. Despite the main focus being on the influence of age on imitation, Nielsen also tested whether female participants imitated more than male participants and found no statistically significant difference.

Babel et al. (2014) reached a similar conclusion to that of Namy et al. (2002). Their experiment consisted of a shadowing task and they also considered gender and the speaker's attractiveness as a factor influencing phonetic imitation. The study focused on vowel quality imitation and also quantified imitation using a perceptual similarity test. Their results showed that women imitated more overall and showed most imitation when shadowing attractive male models. Babel et al. (2014) also addressed the difference between the experiments of Pardo (2006) and Namy et al. (2002). Due to Babel et al. using a similar experimental procedure and obtaining similar results to those of Namy et al., they suggested that the difference in the results of Namy et al. and Pardo may also be caused by the different experiment procedures as Namy et al. used a shadowing task while Pardo's experiment (2006) was based on a conversational task. This hypothesis might also explain why the results of Nielsen (2014) do not entirely match those of other studies.

Pardo et al. (2017) did not find any differences between gender in their experiment. Their experiment featured a shadowing task and focused not only on the imitation of male and female participants but also on how this factor influenced imitation within the relation between the model talker and the participant. However, no significant influence on phonetic imitation was found.

The results overall seem rather conflicting. However, as these researchers suggested, many other factors may have played a role in the results of the aforementioned studies. The results could have been influenced by different social settings (Namy et al., 2002; Pardo et al., 2010) or the different experimental procedures (Babel et al., 2014). Further research is required to reach a clear conclusion about the role of interlocutors' gender as a predictor of phonetic imitation.

3.3 Social Selectivity and Opinion

Social selectivity and the opinion of a listener on the speaker were also studied as a phenomenon which could influence phonetic imitation. Babel focused on this factor in some of her research

papers, she conducted experiments focusing on these factors related to dialect convergence (Babel, 2010), visual prompts (Babel, 2012) and attractiveness of the voice (Babel et al., 2014). From the results of her studies, she mainly hoped to determine whether phonetic imitation can be considered a socially driven phenomenon.

Babel (2010) asked whether vowel imitation would be influenced by model speaker liking. In her experiment, participants from New Zealand took part in baseline recording, a pre-task block, a shadowing task with recordings of an Australian model talker and post-shadowing reading. During the pre-task block, participants were randomly assigned either to a positive condition which would determine whether the participants were given a description that would describe the Australian speaker in a positive light or a negative condition which would do the opposite. The results suggested that participants who viewed the Australian speaker in a positive light were more likely to imitate than those who did not.

In the study of Babel (2012) the participants took part in a shadowing task that examined imitation of vowel quality. Each participant was placed under one of four different conditions in the task. Two model talkers with a Californian accent were recorded, the only difference between the talkers was their race (there was a black and a white talker). Participants were randomly assigned to one of the model talkers and there was a visual prompt of the speaker or not. The participants with no visual prompt were later asked to identify the race of the model talker. Babel found that they were not able to identify the race of the speakers reliably. The social selectivity of the speakers was measured through attractiveness rating due to the high correlation between attractiveness and social selectivity. The results showed that the participants imitated the model talkers selectively. The presence of a visual prompt caused more imitation than the absence of a visual prompt. Furthermore, when met with the present visual prompt condition, both male and female participants imitated the black talker's vowel /æ/, but while females had the same results with the white talker, males imitated the vowel /ɑ/ more with the white talker. For the black talker, no significant relation between attractiveness and imitation was found, however, for the white talker, the males were not likely to imitate him when they rated the model more attractive while females were more likely to imitate the speaker, the more attractive they rated him. Babel concluded that phonetic imitation is 'selective from both a phonetic and social perspective' (2012, p. 188). She concluded that the results were rather mixed though results based on the presence of the visual prompt suggested that social context played a role in phonetic imitation.

In Babel et al. (2014), social selectivity was also studied with regard to voice types and their relation to phonetic imitation. The participants took part in a shadowing task with

recordings from model talkers whose voices were rated using a rating scale as attractive, unattractive, typical or atypical. The study focused on vowel quality imitation. The highest degree of imitation was found for typical and attractive voices. Male participants showed a similar amount of imitation for attractive males, attractive females and not-attractive males. The only exception was the group of not-attractive models to whom the participants imitated less. Female participants imitated more overall, but the greatest degree of imitation was in response to unattractive, atypical female model talkers. They also showed the least imitation of unattractive and atypical male model talkers. Babel et al. (2014) concluded that social preferences influence phonetic imitation and suggested that imitation by women is more likely to be influenced by social preferences.

3.4 Social Engagement

Babel (2012) examined whether the visual prompt of a model talker influences phonetic imitation of vowel quality. As she tried to prove that phonetic imitation is a social phenomenon, Babel predicted that the more social the context would be, the greater the degree of imitation should be. She argued that presenting more social information led to more social engagement which is why in her experiment more social context was ensured through the addition of digital images displayed to the participants. Babel found that the participants imitated the model with a picture significantly more than when the picture was not present. She concluded that the data suggested that phonetic imitation is influenced by a context that is more socially engaging.

Similar results were found by Dias and Rosenblum (2011) who looked for differences between imitation with or without seeing the model talker. The participants took part in a conversational task. They were sorted into pairs where each participant was assigned a space with nine boxes with nine items. They could not see the boxes, but they were allowed to manipulate the items in them by touch. The participants were then asked to match the positions of their items, one at a time, with the item positions of their partners. The results were then assessed using a perceptual similarity task with a separate group of listeners which showed that greater imitation occurred with visual contact between the interlocutors. Dias and Rosenblum (2016) later tried to replicate the results from their 2011 study using a shadowing task. They argued that the environment of the shadowing task would allow more control over the manipulation and presentation of the stimuli. The results matched those of Dias and Rosenblum (2011) which led to their conclusion that even in a noninteractive context, visual information can enhance phonetic imitation. This conclusion matches the conclusion of Babel (2012) who

also found that participants who were shown a visual prompt in the form of a picture imitated more. Overall, these findings suggest that social engagement, at least in the form of a visual prompt, can influence phonetic imitation.

3.5 Conversational Dominance

The role of the speaker in communication was also considered a potential influence of phonetic imitation. Several studies hypothesised that the dominance of a speaker would influence phonetic imitation. Giles (1973) introduced his model of accent mobility and proposed that speakers tend to imitate more dominant speakers. In his experiment, the speakers of a perceived higher status due to them having a Received Pronunciation accent were imitated by other speakers. Because of these results, Giles proposed that speakers use imitation as a means to become more similar to their interlocutor and identify themselves with them.

Pardo (2006) conducted a conversational task experiment which assigned roles to same-sex pairs of speakers who completed tasks. The roles in question were ‘givers’, who were instructed to give instructions and were therefore deemed as the more dominant role, and ‘receivers’, who received and followed the instructions from givers. The shifts in post-exposure productions were quantified in a perceptual similarity task. Based on Giles’s theory (1973), Pardo hypothesised that receivers would imitate givers more than givers would imitate receivers. The overall results, however, showed that givers imitated receivers more. An additional analysis showed that the factor of assigned roles interacted with gender. Male participants showed more imitation of givers, while female participants imitated receivers more. Pardo concluded that these results indicated that the participants may have interpreted dominance in the experiment differently depending on their gender or some other unknown factor.

Pardo et al. (2010) elaborated on the study by Pardo (2006), asking whether the findings would be replicated while explicitly instructing the participants to imitate. They examined the imitation of articulation rates and vowel formants via acoustic measurements. Pardo et al. (2010) hypothesised that explicit instruction would lead to greater levels of imitation. If this hypothesis was confirmed in their experiment, they reasoned that imitation should not be completely assigned to the influence of gender, but that these factors might also be connected to the participants’ attention. The results of Pardo et al. (2010) indeed replicated the results of Pardo (2006). However, when the givers were instructed to imitate, only male givers imitated while female givers diverged from the model talker. When the receivers were instructed to

imitate, all talkers imitated. The authors concluded that despite the inconclusive results, the speaker role should be considered a significant factor influencing phonetic imitation and more research would be needed to understand how it interacts with other factors.

Pardo et al. (2013a) further elaborated on the findings of Pardo (2006) and Pardo et al. (2010). In their study, Pardo et al. (2013a) tried to simulate a more natural situation by alternating the participants' roles during a conversational task and seeing under what conditions imitation would occur. Participants took part in a map task where they switched roles five times. The results showed that despite the role-switching, participants who were originally givers kept their dominant role in conversation despite the alternation. The dominance of the participant was determined by the total time the participant spent talking. The participant that talked more was considered the more dominant speaker in the conversation. Overall, the participants imitated when the roles were switched and not in their original roles. However, due to the alternation, former givers, now receivers, imitated givers, as opposed to the previous studies where receivers did not really imitate. Pardo et al. proposed that this was due to an underlying process which was induced by role-switching. They conclude that the differences in results of experiments concerned with conversational dominance may happen because of differences in tasks. In summary, the effect of speaker role or dominance in the communicative situation appears to be important, but a better understanding of how it interacts with other factors is needed before reliable predictions of imitation probability, direction and degree can be made.

3.6 Task Engagement

Researchers also asked whether task engagement could be a factor influencing phonetic imitation. Biro, Toscano, and Viswanathan (2022) introduced task engagement as a factor influencing behaviour and attention that has not yet been thought of as a factor of phonetic imitation. Participants took part in collaborative tasks with either high or low engagement levels. Their hypothesis that high-engagement tasks would lead to more phonetic imitation overall was confirmed although there were specific conditions in which the imitation occurred. VOT imitation was found only in the high engagement level and only in tokens starting with voiced consonants, vowel duration for tokens starting with voiceless consonants was also imitated in the high engagement level. The authors concluded that phonetic imitation is likely influenced by task engagement. However, they also suggested that their sample size might have been too small to see consistent effects and that in the higher engagement task, speakers talked more overall which might have affected the speakers' imitation due to more exposure.

3.7 Exposure

Goldinger (1998) examined whether repeated exposure would have influence on phonetic imitation in a shadowing task. Goldinger tried to find evidence for the exemplar view of lexical representation, which predicts that the tokens with the highest exposure should be the most imitated. The stimuli were divided into four categories depending on the number of repetitions of the stimulus before shadowing (0, 2, 6 or 12 repetitions) and the experiment consisted of several listening and shadowing tasks which were alternated with each other so that each participant would go through all conditions of repetitions. The resulting productions were then used in a perceptual similarity task in order to quantify imitation. Goldinger's hypothesis was confirmed by the results of the experiment.

Goldinger and Azuma (2004) later tested the hypothesis of Goldinger (1998) in an experiment which consisted of four parts and aimed to study exposure over the course of two weeks. The four parts consisted of a baseline recording, a training stage during which the participants listened to recordings, a re-recording stage which took place a week after the training stage and which repeated the procedure of the baseline recording, and a recognition task during which the participants were asked to select the original words in a mix of words with matched word frequency. The number of repetitions was manipulated in the training stage. The results matched those of Goldinger (1998): Goldinger and Azuma (2004) found that imitation was greater after more repetitions.

Nielsen (2014) attempted to replicate Goldinger's results (1998) using only two kinds of stimuli regarding exposure: the words that participants were exposed to in an exposure phase and the words they were not exposed to in an exposure phase. However, her results did not show any imitation tied to exposure. She concluded that her results did not directly oppose the exemplar-based theories (due to the results of the factor of age following the exemplar-based hypothesis, see 3.1), but rather led to the assumption that the effect of exposure was subtle.

3.8 Word Frequency

Goldinger (1998) also investigated imitation of acoustic patterns (e.g., fundamental frequency and word duration) in words and nonwords in relation to word frequency and exposure of the participants to the stimuli using a shadowing task. He found that words which occur with high frequency were often imitated less than low-frequency words. Goldinger believed that this is because '[high frequency] words inspire 'abstract' echoes' (1998, p. 255), meaning that during perception, the number of activated traces of high-frequency words would be so high that the

details which are contained in the traces would become obscured and imitation would therefore be less likely. Similar results were later replicated in the experiment by Goldinger and Azuma (2004). Nielsen (2011) also managed to replicate the frequency effect in a non-shadowing experiment investigating the imitation of VOT.

On the other hand, Pardo et al. (2013b) did not find an effect of lexical frequency. Pardo et al. conducted two experiments, the second of which was based on a shadowing task and investigated the influence of word frequency and neighbourhood density of the target words. While the authors found imitation, their results did not show a reliable word frequency effect on phonetic imitation. Pardo et al. suggested that the failure to reach similar findings as the previous studies could be caused by familiarity with the speaker's voice as research found that familiarity with the speaker's voice might negatively affect the effect of word frequency (Bradlow and Pisoni, 1999).

Dias and Rosenblum (2016) conducted a shadowing experiment concerned mainly with the influence of the presence of visual prompts, but also with word frequency and phonological neighbourhood density. Their results showed that low-frequency words were imitated more than high-frequency words and they concluded that the results suggested that imitation is influenced by word frequency. They also addressed the discrepancy in results between their study and Pardo et al. (2013b) and suggested that Pardo et al. used monosyllabic stimuli which could have influenced the results.

The discrepancy in the results between Pardo et al. (2013b) and the other studies cited was later addressed by Pardo et al. (2017). They mentioned that Nielsen (2011) and Dias and Rosenblum (2016) used one model talker only and that the shadowers in Dias and Rosenblum were mostly female while in Pardo et al. (2013b) there were 20 model talkers and the participants consisted of an equal number of same-sex pairs. The 2017 study by Pardo et al. attempted to replicate Goldinger's (1998) results once more with a significantly higher number of participants. The results were similar to those of Pardo et al. (2013b) and they showed that phonetic imitation was not significantly influenced by word frequency. Pardo et al. (2017) also found that female participants imitated words with low frequency more while it had no significant effect on male participants. They mentioned that '[t]hese interaction effects help explain some of the inconsistencies observed across the literature with respect to talker sex and word frequency' (Pardo et al., 2017, p. 646) as the participants in the previous studies where word frequency was found to be a significant factor of imitation were mostly female. Pardo et al. (2017) also found that disyllabic words are imitated more than mono-syllabic (see 3.10).

Because of this information, Pardo et al. suggested that the results of Dias and Rosenblum (2016) were influenced by the disyllabic item set which Dias and Rosenblum used.

3.9 Phonological Neighbourhood Density

Phonological neighbourhood density has been also considered a factor of phonetic imitation. Phonological neighbourhood density of a word is equal to the number of words which are different in one phoneme from the word whose density is being measured (Pardo et al., 2013b). This factor was studied along with word frequency (Pardo et al., 2013b; Dias and Rosenblum, 2016).

Pardo et al. (2013b) investigated neighbourhood density in relation to phonetic imitation through an experiment based on a shadowing task. They used the terms ‘Easy’ words which describe words with low neighbourhood density and high word frequency and ‘Hard’ words which have high neighbourhood density and low frequency. This terminology was derived from the fact that Hard words are perceptually the most difficult (Pardo et al., 2013b). Pardo et al. predicted that Hard words would be more imitated than Easy words. However, their results did not find a statistically significant impact of phonological neighbourhood density on phonetic imitation. As was mentioned in 3.8, Pardo et al. proposed that familiarity with the stimuli could have influenced their results.

Dias and Rosenblum (2016) on the other hand predicted that low neighbourhood density words would be more imitated in their shadowing experiments. They found that imitation was indeed greater for words with fewer phonological competitors in their lexical neighbourhood. As Pardo et al. (2013b) did not find reliable influence of phonological neighbourhood density perhaps the effect of word frequency cancelled out the effect of neighbourhood density. Dias and Rosenblum (2016) concluded that their data suggested that neighbourhood density influences imitation. As mentioned in 3.8, the authors also addressed the discrepancy in the results of Pardo et al. (2013b) when compared with the other studies and they proposed that the unexpected results of Pardo et al. could be influenced by the stimuli used in the study of Pardo et al. as it was monosyllabic.

3.10 Number of Syllables in Words

Pardo et al. (2017) asked whether the number of syllables of the stimuli would affect phonetic imitation. They claimed that a higher number of syllables could lead to more imitation as longer words would offer more chances to imitate (p. 643). Goldinger’s disyllabic item set (1998) was

used in comparison to the monosyllabic item set used in Pardo et al. (2013b) in a shadowing task. Pardo et al. (2017) found that phonetic imitation was significantly greater in the disyllabic item set when compared to the monosyllabic item set. This seems to support the idea of Dias and Rosenblum (2016) to a certain degree as they proposed that the results of Pardo et al. (2013b) were different due to the stimuli being monosyllabic. However, Pardo et al. (2017) concluded that the greater imitation was caused by the factor of the number of syllables and that word frequency had no significant effect on phonetic imitation while Dias and Rosenblum (2016) concluded that word frequency influenced the imitation in their experiment.

3.11 Language Distance

Language distance is another factor that was studied as an influence on phonetic imitation. Kim et al. (2011) investigated whether same-dialect pairs would show more imitation than different-dialect pairs. Using an experiment based on a picture description and completion task, a conversational task encouraging balanced roles within a conversation, Kim et al. examined native speakers of American English, Korean and Chinese. The speakers were divided into pairs who spoke the same dialect of the same native language, a different dialect but the same native language and a different native language. The same dialect pairs consisted of participants who were either Korean speakers speaking the same dialect or English speakers speaking the same dialect. In the pairs where participants spoke different dialects of the same native language, participants were either Korean speakers with different dialects or English speakers with different dialects. The pairs with different native language consisted of an English speaker who was paired with either a Korean or a Chinese speaker. All conversations between the participants were in English except when both participants spoke the same native language which was not English. The authors found that while the pairs speaking different dialects and different native language had similar results, the pairs with the same dialect imitated reliably more. Although the data might suggest that language distance could be a factor influencing phonetic imitation, Kim et al. mentioned that some factors of the experiment, such as the possibility of speakers spontaneously choosing a more dominant/submissive role in the conversation (see 3.5), could have also influenced the results.

Paquette-Smith et al. (2022) conducted a study where they expected to find similar results to Kim et al. (2011) regarding the influence of language distance on the imitation of manipulated VOT. In their experiment, child and adult speakers of Canadian English listened to three model talkers: a speaker of Canadian English, a speaker of Australian English and a

non-native Mandarin talker who spoke English. The results indicated that the imitation of the Canadian English speaker was not significantly different from the imitation of the Australian English speaker. However, they found that the participants imitated the Mandarin-accented speech the more than the Canadian English speaker. Paquette-Smith et al. concluded that no clear reason exists for the difference in the results between them and Kim et al. (2011), although there was the possibility that social selectivity influenced imitation in the experiment by Paquette-Smith and colleagues.

Tobin (2022) tested the role of linguistic background in the imitation of VOT using the phonetic distance of languages. In his shadowing experiment, he investigated whether the speaker's native language would influence VOT imitation. The native languages of the speakers in this experiment were English, Korean, and Spanish. Tobin's hypothesis that VOT imitation would depend on the native language of the speakers was confirmed by the results. While all groups showed imitation to English stimuli, the Korean speakers showed the greatest imitation and the Spanish speakers the lowest. Tobin reasoned that the outcome was caused by 'the relative stability of the language-specific laryngeal-oral coordination patterns that give rise to VOT' (2022, p. 11) meaning that the overall articulatory stability of the VOT productions in the language influenced the outcome of this experiment.

3.12 Contrast Maintenance and Perceptual Salience

Salience was defined in phonology as a feature that is 'more visible or noticeable both for the linguist and the language user' (Rácz, 2013, p. 23). Honeybone and Watson add that a feature becomes more salient when 'involved in the expression of a phonological contrast' (2013, p. 308). Perceptual salience was suggested as a factor influencing phonetic imitation in 2015 by Podlipský and Šimáčková. Podlipský and Šimáčková studied the preservation of phonological contrast on account of a disagreement they found between the studies of Nielsen (2011) and Mitterer and Ernestus (2008). In her paper, the results of Nielsen (2011) showed that the imitation of VOT of voiceless stops in her experiment was asymmetrical. While the participants imitated the extended VOT, they did not produce the reduced VOT. Nielsen (2011) proposed that one of the possible reasons for the asymmetrical pattern of imitation found is that speakers tried to maintain the contrast between phonological categories as the reduced VOT production of voiceless stops would threaten the contrast between the voiceless stops and the voiced stops and the extended VOT would not have such effect. However, Podlipský and Šimáčková (2015) mentioned that these results contradicted the findings of Mitterer and Ernestus (2008) who

reported that Dutch speakers did not imitate the lengthened VOT, but did imitate the shortened VOT.

In order to test Nielsen's hypothesis, Podlipský and Šimáčková conducted an experiment with a shadowing task where they examined the implicit imitation of Czech speakers of /d/-prevoicing, as Czech contrasts pre-voiced stops with unaspirated stops, and imitation of a duration of the vowel /u:/, as Czech contrasts long and short vowels. The participants imitated the prolonged /d/-prevoicing condition significantly and did not imitate the shortened /d/-prevoicing. However, for the imitation of vowel duration, while the participants imitated the prolonged /u:/ duration in the shadowing session with the prolonged condition and both post-shadowing readings, they also imitated the reduced /u:/ duration in the shadowing session. An independent test which followed found that 'the reduction of /d/-prevoicing is less salient to Czech listeners than its extension' (Podlipský and Šimáčková, 2015, p. 4). Based on the results, Podlipský and Šimáčková concluded that the contrast preservation hypothesis does not completely prevent phonetic imitation. Their conclusion was later addressed by Kim and Clayards (2019) who proposed that perceived speech tempo could have influenced the results.

Kim and Clayards (2019) asked whether perceptual salience or contrast preservation can influence phonetic imitation. They studied vowel duration and spectral contrast between the vowels /ɛ/ and /æ/ as these vowels are 'some of the most regionally variable vowels in North America due to sound change such as the Canadian Shift' (Boberg, 2008, as cited in Kim and Clayards, 2019, p. 772), therefore their manipulation would remain acceptable to the participants. Their stimuli were divided into unambiguous vowels and ambiguous vowels. The ambiguous vowels were selected by native speakers who were instructed to choose the most ambiguous option of a vowel spectrum continuum created between /ɛ/ and /æ/. The duration of the vowels was also manipulated into extended and reduced durations. Kim and Clayards argued that reducing the duration of /æ/ would threaten the phonological contrast with /ɛ/, and similarly, extending the duration of /ɛ/ would threaten contrast with /æ/. The study consisted of a baseline task and an imitation task where the participants were explicitly instructed to imitate the stimuli. The results showed that unambiguous vowels were imitated significantly more. Moreover, the imitation increased with longer vowel duration irrespective of vowel quality. The authors proposed that this happened because longer vowels 'provide better information about the target spectral quality than shorter vowels' (p. 779). Furthermore, they found that imitation of perceptually salient vowel durations (prolonged and reduced) was significantly greater than the imitation of the rest of the vowel durations. One of the possible explanations Kim and

Clayards put forward is that speakers tend to try and preserve phonological contrast, but that vowel duration is not as important to maintain the contrast as other phonetic properties such as vowel quality for this particular contrast in question. The potential influence of contrast maintenance on imitation and the exact role of the perceptual salience of the imitated pronunciation feature therefore remain unclear.

4 The Present Study

As can be seen in the literature reviewed above, it is evident that while phonetic imitation is a frequently attested phenomenon which occurs even in brief exchanges between interlocutors there are a number of influences it is subject to. The current state of research does not allow a conclusive evaluation of how each of those factors influences the probability of occurrence, direction and the extent of imitation. Furthermore, the review illustrates that there are probably quite complex interactions between the different factors (e.g., between participants' gender and role in conversation or even gender and property of the stimulus at hand, namely word frequency). Therefore, there is need to explore the phenomenon of phonetic imitation further and throw more light at whether and how different factors influence imitation, which is why the present study investigates subconscious imitation of prolonged or reduced vowel duration previously observed by Podlipský and Šimáčková (2015) and whether it could be replicated in the presence of a short introductory carrier phrase with a fixed speaking tempo. This chapter describes the experiment which was conducted as part of this thesis, the methodology used and the results.

The procedure of the experiment replicated the procedure of Podlipský and Šimáčková (2015) who examined whether speakers tend to limit phonetic imitation in order to preserve phonological contrast. The present study specifically focuses on their experiment concerning /u:/ duration imitation and elaborates on their findings.

4.1 Methods

In this chapter, I describe the methods used in the present study consisting of the procedure of the experiment, the stimuli chosen for the experiment and the participants of the experiment. The experiment consisted of two sessions which included a baseline reading task which was present only in the first session, a shadowing task and a post-shadowing reading (included in both session 1 and session 2). The experiment design replicated that of Podlipský and Šimáčková (2015) who examined whether perceptual salience and the need to preserve

phonological contrasts influence phonetic imitation, with the exception that in their study, Podlipský and Šimáčková did not use short carrier phrases during shadowing. The participants were exposed to prolonged and reduced duration of the phonologically long vowel /u:/ in the shadowing sessions.

4.1.1 Material

The material consisted of 49 Czech disyllabic words and 5 short carrier phrases. Nineteen of the stimulus words contained the vowel /u:/ (see 7.1) and 30 were filler words. The words were part of the stimulus material used in Podlipský and Šimáčková (2015). A female native Czech speaker was recorded as the model talker for the stimuli in Podlipský and Šimáčková (2015) and another female native Czech speaker who sounded very similar to the previous model talker was additionally recorded for carrier phrases. The stimuli were manipulated to receive a prolonged /u:/ condition which was extended in duration by a factor of 1.277 or a reduced /u:/ condition which was reduced by a factor of 0.763. These manipulations were chosen by Podlipský and Šimáčková (2015) as a magnitude that would be unnoticed by native Czech speakers.

The words were preceded by short carrier phrases (see 7.2). The carrier phrases were added to contextualize the speaking tempo as proposed by Kim and Clayards (2019) who suggested that the results of Podlipský and Šimáčková (2015), where both extended and reduced /u:/ duration were imitated, could be caused by perception differences in the speech tempo. Therefore, in order to test whether speakers imitate without the possible influence of these perception differences, this experiment added the short phrases to the stimuli used in Podlipský and Šimáčková's experiment. Each of the carrier phrases consisted of five syllables and had a fixed speaking tempo. The carrier phrases were recorded multiple times and the final version of each phrase was selected and minimally edited in Praat (Boersma, Weenink, 2022) ensuring that each had the exact same duration of 1 s. The carrier phrases only preceded the target words and did not embed them in order to sound natural to native Czech listeners.

4.1.2 Procedure

The experiment took place in a sound-insulated booth and the participants were recorded individually using a Zoom H4n digital recorder. All information and stimuli were presented using a Praat Demo window visible on a PC laptop monitor that was situated on a table in front of the participants. The procedure of the experiment replicated the procedure of Podlipský and Šimáčková (2015). Participants attended two sessions, the first of which was divided into three

parts and the second into two parts. The two sessions were separated by at least 24 hours. Each of the parts in the session was introduced by a set of instructions and information about the session part. Within each session, participants were exposed to either extended or reduced durations of /u:/. Participants wore headphones throughout the entirety of the experiment in order to prevent differences in the recordings caused by the Lombard effect (i.e., more careful pronunciation when wearing headphones as compared to not wearing them).

The first session consisted of a baseline reading task, a shadowing task and a post-shadowing reading task (identical in form to the baseline). At the start of the first session, a list of all stimuli and fillers was presented to the participants. The second session consisted of a shadowing task and a post-shadowing reading (i.e., no pre-shadowing task was repeated). In the baseline and the post-shadowing reading task, the stimuli were presented one by one on the screen automatically. In the shadowing sessions, the stimuli were played through headphones after the short carrier phrase. The participants were instructed to repeat only the words, not the carrier phrases. In all sessions, there was a fixed 2.7s interval between the stimuli whether presented visually as text or auditorily. The sessions were separated by breaks and participants could pause and resume at any time during the experiment.

In each of the shadowing tasks, the participants received recordings with different conditions (reduced or prolonged vowel duration). The participants were assigned into groups with different conditions randomly. As presented in

Table 1: Conditions across groups and sessions, the first group was assigned the reduced vowel duration in the first session and the extended vowel duration in the second session. The second group received the extended condition in the first session and the reduced condition in the second session.

Group	Session	Manipulation
1	1	reduced
2	1	extended
1	2	extended
2	2	reduced

Table 1: Conditions across groups and sessions.

4.1.3 Participants

Seventeen native Czech speakers aged between 18 and 22 who were all students took part in this experiment. All speakers were female in order to allow comparison to Podlipský and Šimáčková (2015) who also used exclusively female speakers and also in order to rule out the possible influence of gender on the results. Most of the participants came from the Olomouc region (the rest came from the Prague region, the Brno region, the Plzeň region and the Ústí nad Labem region). None of them originated from or lived in the Ostrava region due to the possibility of the speakers shortening the long vowels. All participants reported having normal hearing.

4.2 Results

The elicited productions were annotated in Praat v. 6.3.03 (Boersma, Weenink, 2022) for word beginnings and ends and /u:/ boundaries according to the principles defined by Machač and Skarnitzl (2009).

Out of the total of 1615 recorded tokens, 1558 tokens could be used for the analysis. The rest (forming 3.5% of the recordings) were not used as the word beginnings or ends were cut off due to participants missing period within which sound was recorded automatically within each shadowing trial.

First, extreme values were checked for errors of measurement (none had occurred). The resulting raw data is presented in

Figure 1 displaying the distributions of durations of /u:/ measured in seconds across the different conditions.

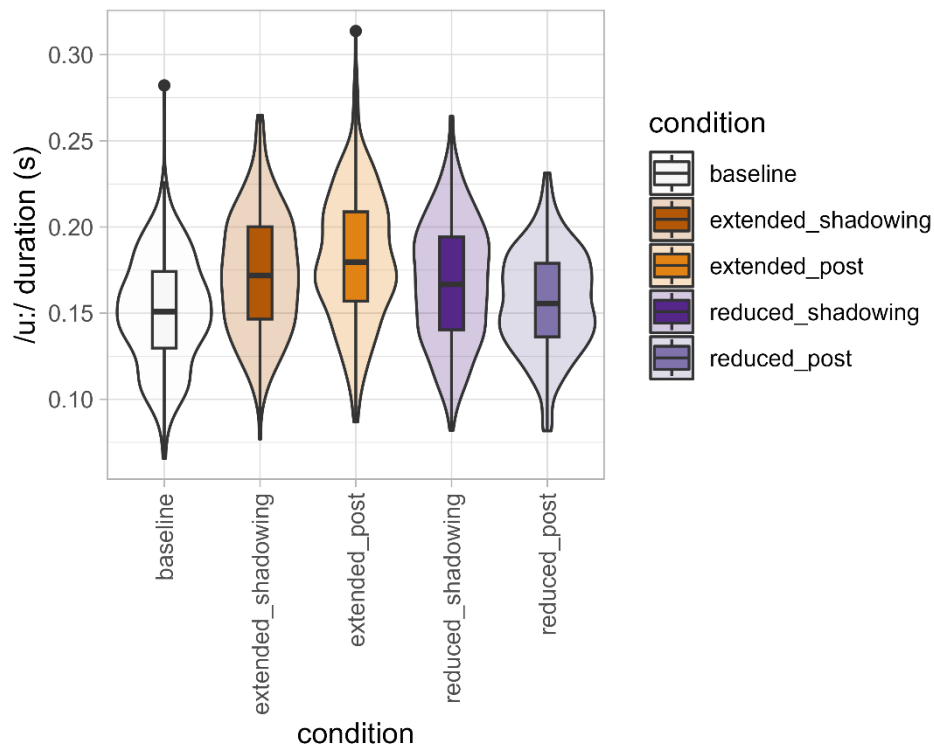


Figure 1: Durations of /u:/ across conditions visualised in a violin plot.

Although the shadowing tasks used the carrier phrases designated to ensure that differences in the perception of the speech tempo between the two shadowing conditions (extended /u:/ and reduced /u:/) were minimized, the /u:/ duration data were also normalised by expressing the duration of the long vowel /u:/ as a percentage of the duration of the word in which it was uttered. This was possible since the exact same set of target stimulus words was used in each condition. Apart from eliminating local variation of speech tempo as a factor, the normalisation was performed so that the resulting values could be compared directly with the values of Podlipský and Šimáčková (2015). The normalised data are shown in Figure 2.

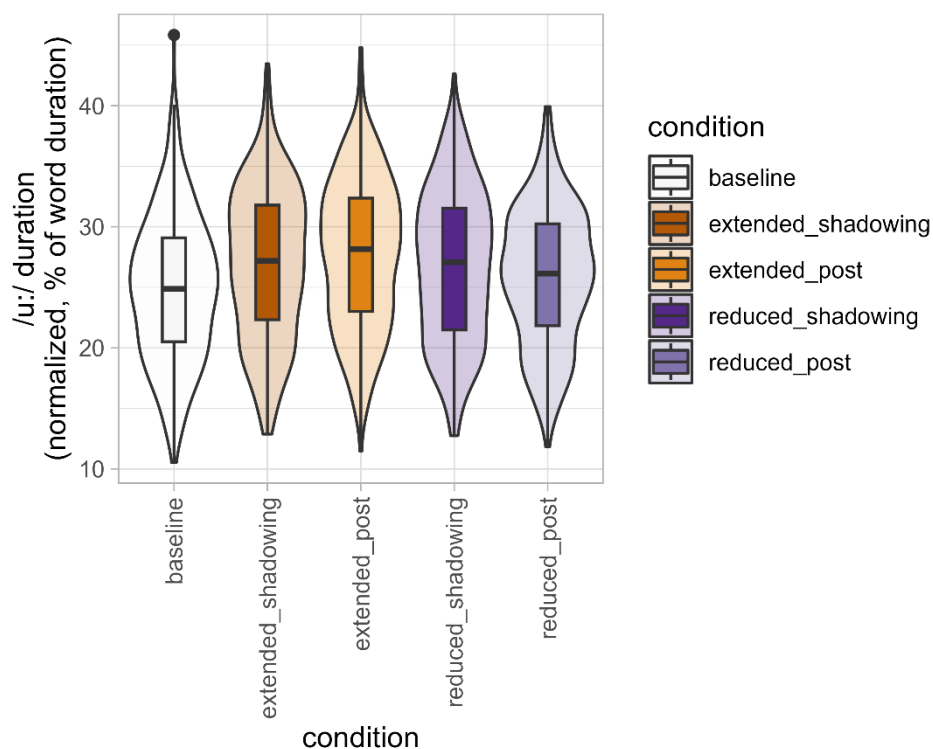


Figure 2: Normalised /u:/ durations (expressed as percentage of word duration) across the five conditions.

The data in

Figure 2 seem somewhat more dispersed as compared to

Figure 1, which is not surprising, as the words were of different durations. While all words were disyllabic, some consisted of more segments than others, such as *schüdnost* /sxu:dnost/ with eight segments versus *fúze* /fu:ze/ with four segments.

The raw distributions of durations, as seen in the violin plots both in Figure 1 and Figure 2, seem to suggest that the shadowing productions, as well as the post-shadowing reading, after exposure to extended /u:/ durations led to somewhat longer /u:/ durations than those in the baseline reading. At the same time, the raw distributions do not point to an in-shadowing or a post-shadowing shortening of /u:/ duration after exposure to reduced /u:/ duration in the manipulated stimuli.

To test this statistically, a mixed linear regression model was fitted to the normalised /u:/ duration data. The analyses were performed (and all figures were drawn) using R (R Core Team, 2022; Bates et al., 2015; Singmann et al., 2022; Wickham, 2016). The fixed effect was condition with five levels: the baseline reading task, the shadowing of extended /u:/ words (‘extended shadowing’), the reading task after the shadowing of extended /u:/ words (‘extended post’), the shadowing of reduced /u:/ words (‘reduced shadowing’) and the reading task after

the shadowing of reduced /u:/ words ('reduced post'). This predictor was treatment-coded with the baseline level serving as the reference level. The performance in the shadowing tasks and the post-test were the treatment conditions. The model also included two random effects: the random effect of participant with varying intercepts and slopes for condition and the random effect of word with varying intercepts.

Table 2 presents the resulting regression coefficient estimates.

Table 3 shows the model predicted values for each of the conditions and Figure 3 plots these values. The intercept, corresponding to the /u:/ durations in the baseline reading task, as was mentioned above, is estimated to be about a quarter of the word duration. When the participants were shadowing words with the extended durations of /u:/, the /u:/s they produced were around 2% of word duration longer than they were in the baseline productions (estimate = 2.09 p. p., SE = 0.62, t = 3.39, p < .01). When speaking in the post-test after being exposed to the extended /u:/ durations in the shadowing session, their productions of /u:/ were reliably longer (estimate = 2.54 p. p., SE = 0.46, t = 5.49, p < .001). For the shadowing session where participants shadowed words with reduced /u:/ durations, the /u:/ productions were also significantly longer than baseline /u:/s (estimate = 1.89 p. p., SE = 0.45, t = 4.21, p < .01). On the contrary, the post-test after being exposed to reduced /u:/ durations did not have durations that would be reliably different from the baseline values (p > .05).

Condition	Estimate	Std. Error	df	t value	Pr(> t)
Intercept/Baseline	24.7741	1.3456	23.5242	18.4113	0
Extended Shadowing	2.0878	0.6153	15.9393	3.3934	0.0037
Extended Post	2.5446	0.4636	16.1187	5.4893	0
Reduced Shadowing	1.8883	0.4489	15.8544	4.2067	7e-04
Reduced Post	0.5701	0.4214	16.046	1.3529	0.1948

Table 2: Coefficient estimates across conditions.

Condition	Predicted	Std. Error	Conf. low	Conf. high
Baseline	24.7741	1.3456	22.1347	27.4135
Extended Shadowing	26.8619	1.3552	24.2037	29.5201
Extended Post	27.3187	1.3482	24.6741	29.9632
Reduced Shadowing	26.6623	1.3629	23.9890	29.3356
Reduced Post	25.3442	1.3103	22.7740	27.9144

Table 3: Model-predicted values for the different conditions.

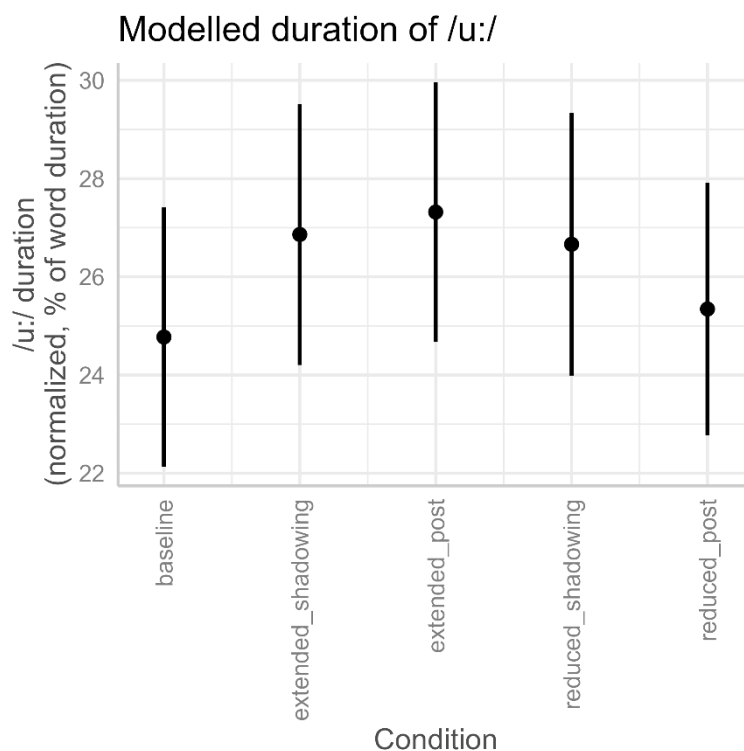


Figure 3: Model-fitted /u:/ duration across conditions.

5 Discussion

This thesis reviews the phenomenon of phonetic imitation and its relation to the factors that influence its probability, direction and degree. The main goal of this thesis was to elaborate on the study of Podlipský and Šimáčková (2015) and specifically see whether and how maintenance of the phonological contrast affects phonetic imitation.

In order to elaborate on Podlipský and Šimáčková (2015), the experiment conducted as part of the present thesis replicated that of Podlipský and Šimáčková and specifically asked whether native Czech speakers extend the duration of long vowel /u:/ after being exposed to extended /u:/ and reduce the duration of this vowel after being exposed to reduced /u:/ even if the stimuli follow a five-syllable carrier phrases with a constant speaking tempo in a shadowing session as was suggested by Kim and Clayards (2019).

Like in the previously mentioned studies, in the current experiment, participants imitated the model talker (e.g., Kim and Clayards, 2019; Nielsen, 2011; Nielsen, 2014; Podlipský and Šimáčková, 2015). The results showed that when compared to the baseline productions, participants prolonged the /u:/ durations in both shadowing tasks and the post-test reading task after being exposed to the extended /u:/ durations although they did not imitate reduced /u:/ in neither the shadowing task with the reduced /u:/ condition nor the post-task after

being exposed to the reduced /u:/ durations. Overall, these results suggest that the hypothesis of Nielsen (2011) and Podlipský and Šimáčková (2015) was correct and that speakers strive to maintain phonological contrast between neighbouring speech and sound categories when it is threatened and that the effort to maintain the contrast affects the direction and degree of phonetic imitation. Despite the carrier phrases with a fixed speaking tempo, participants in this experiment did indeed imitate only the conditions that enhanced phonological contrast between the categories of long vowel /u:/ and the short vowel /u/ which leads to the conclusion that contrast maintenance does seem to influence phonetic imitation.

When compared to the results of Podlipský and Šimáčková (2015), the results reported in this thesis match partly. Like in the experiment of Podlipský and Šimáčková, longer durations of /u:/ were produced by the participants in the current thesis in the shadowing task and the post-shadowing reading after the exposure to the extended /u:/ than in the baseline reading task. However, results in the reduced duration conditions differed from those of Podlipský and Šimáčková (2015). The values from the post-test reading task following the shadowing of reduced /u:/ were not significantly different from the baseline values which does not differ greatly from the results of Podlipský and Šimáčková in the same condition. The main difference between the results found in the present experiment and those reported by Podlipský and Šimáčková was found in the shadowing session when the participants were exposed to the reduced /u:/ durations.

While Podlipský and Šimáčková (2015) found that speakers reduced the /u:/ durations, the results of the current experiment show that participants do not imitate reduced /u:/ durations, instead the /u:/s were slightly prolonged even after exposure to the words with reduced /u:/ durations. One possible explanation for these results is that speakers are not ready to imitate the shortening of a long category because they, whether subconsciously or not, strive to maintain contrast with a short /u/ category. This would also explain why they would imitate the prolonged /u:/ durations as this does not harm but instead enhances the contrast with the Czech short vowel /u/.

This conclusion also matches the conclusion of Kim and Clayards (2019), as Kim and Clayards concluded that speakers prefer to maintain the phonological contrast between categories because their results showed that speakers imitate phonologically unambiguous vowels more than ambiguous vowels. Although the current results differ in the imitation of vowel duration which was not found to be imitated in line with the hypothesis of Nielsen (2011) and Podlipský and Šimáčková (2015), Kim and Clayards did mention that their results may have been influenced by the fact that the vowels they chose for imitation were of different vowel

quality. While this could have caused the imitation of only prolonged vowel duration in their experiment despite the fact that this should have threatened the phonological contrast, this was not the case for the current experiment as the vowels used in the current experiment did not differ in vowel quality.

A suggestion that may also arise as a hypothesis for the results of this experiment is that the carrier phrases could have influenced the phonetic imitation of the participants. However, during the shadowing sessions, the effects on the /u:/ duration were quite similar. Even if the carrier phrases slowed down the speech tempo and therefore extended the /u:/ durations, word durations were accounted for in the process of normalisation of the data. The speech tempo, therefore, would not have been able to influence the results as the /u:/ durations were analysed as a percentage of the word duration.

6 Conclusion

The aim of this thesis was to review literature regarding possible factors influencing the phenomenon of phonetic imitation and to elaborate on a study by Podlipský and Šimáčková (2015) and to find out whether people imitate reduced or prolonged vowel /u:/. The experiment reported in this thesis found that people extend /u:/ durations even when the stimuli follow carrier phrases in shadowing sessions.

The present results found in the experiment replicate those of Podlipský and Šimáčková (2015) for the extended shadowing and the extended post-shadowing. However, the shortening of /u:/ after exposure to reduced duration of /u:/ was not replicated. These results could potentially suggest that contrast preservation can influence phonetic imitation. In Podlipský and Šimáčková (2015), the degree of imitation-induced shifts was greater after exposure to extended durations than the shifts due to exposure to reduced durations, which were smaller. Perhaps the contrast in Podlipský and Šimáčková (2015) was also protected in this way, but not to an extent which would completely block the phonetic imitation of the reduced long vowel /u:/ but only in a way which attenuated the imitation of reduced /u:/. The current results suggest that in the experiment reported in this thesis, the imitation was blocked completely. There is no obvious explanation as to why the /u:/ durations were extended in the shadowing session where participants were exposed to the reduced durations of /u:/. Perhaps future research is necessary to find out what influenced the imitation in the shadowing session with the reduced condition. As for the possibility that the /u:/ proportions were extended because of the presence of the carrier phrases, the same results were found in Podlipský and Šimáčková (2015) for the extended condition and with the same stimuli, where the carrier phrases were not present which would contradict this suggestion. Moreover, the extended productions were also found in post-shadowing reading following the shadowing of the extended /u:/ durations where the carrier phrases were not present and the task was thus identical to the post-test reading task in Podlipský and Šimáčková (2015). The explanation that seems to be the most logical as no other potential influence suggests itself, is that the participants did imitate the extended /u:/ durations in the extended conditions. This would lead to the conclusion that the absence of imitation of reduced /u:/ durations is indeed ascribable to the speakers' probably subconscious need to preserve the contrast with short Czech /u/.

7 Appendices

7.1 Appendix 1: List of target words with their English translations

Target word	Translation
čůza	<i>mildly vulgar expression used for a woman</i> (N)
fůze	fusion
fůzní	fusional
chůdy	stilts
chůze	walking (N)
schůdek	step
schůdnost	accessibility, <i>in terms of walking</i> (N)
schůdný	walkable
schůze	meeting (N)
Súdán	Sudan
úděs	terror
úsměv	smile (N)
úsvit	dawn
úzce	narrowly
úžas	amazement
úžeh	sunstroke
zůstat	stay (V)
zúžit	narrow (V)
žůžo	<i>used to express excitement or happiness</i> (Adj or Adv)

Note: Czech uses two letters to signify the long /u:/ vowel. While both *ú* and *ů* are used, they signify the same sound for Czech speakers.

7.2 Appendix 2: List of carrier phrases with their English translations

Carrier phrase	Translation
Další slovo je	The next word is
Nyní řekněte	Now say (<i>formal</i>)
Ted' je na řadě	Now it's turn to play
Ted' přehrajeme	Now we will play
Vyslovte slovo	Say the word (<i>formal</i>)

8 Works Cited

- Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, 39(4), 437-456. <https://doi.org/10.1017/S0047404510000400>.
- Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, 40(1), 177–189. <https://doi.org/10.1016/j.wocn.2011.09.001>.
- Babel, M., McAuliffe, M., & Haber, G. (2013). Can mergers-in-progress be unmerged in speech accommodation? *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00653>.
- Babel, M., McGuire, G., Walters, S., & Nicholls, A. (2014). Novelty and social preference in phonetic accommodation. *Laboratory Phonology*, 5(1). <https://doi.org/10.1515/lp-2014-0006>.
- Bates, D., Maechler, M., Bolker, B., Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Softw.* 67, 1–48.
- Biro, T., Toscano, J. C., & Viswanathan, N. (2022). The influence of task engagement on phonetic convergence. *Speech Communication*, 138, 50–66. <https://doi.org/10.1016/j.specom.2022.02.002>.
- Boberg, C. (2008). Regional phonetic differentiation in standard Canadian English. *Journal of English Linguistics*, 36(2), 129-154.
- Boersma, P. & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program]. Version 6.3.03, retrieved 17 December 2022 from <http://www.praat.org/>.
- Bourhis, R. Y. & Giles, H. (1977). The Language of Intergroup Distinctiveness. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 119–135). London, UK: Academic Press.
- Bradlow, A. R., & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener-, and item-related factors. *The Journal of the Acoustical Society of America*, 106(4), 2074–2085. <https://doi.org/10.1121/1.427952>.
- Byrne, D. E. (1971). *The attraction paradigm*. New York, Academic Press.
- Dias, J. W., & Rosenblum, L. D. (2011). Visual Influences on Interactive Speech Alignment. *Perception*, 40(12), 1457–1466. <https://doi.org/10.1068/p7071>.
- Dias, J. W., & Rosenblum, L. D. (2016). Visibility of speech articulation enhances auditory phonetic convergence. *Attention, Perception, & Psychophysics*, 78(1), 317–333. <https://doi.org/10.3758/s13414-015-0982-6>.

- Dragojevic M., Gasiorek J. & Giles H. (2016). Accommodative Strategies as Core of the Theory. In H. Giles (Eds.), *Communication Accommodation Theory: Negotiating Personal Relationships and Social Identities Across Contexts* (pp. 36-59). Cambridge, United Kingdom: Cambridge University Press.
- Giles, H. (1973). Accent Mobility: A Model and Some Data. *Anthropological Linguistics*, 15(2), 87–105. <http://www.jstor.org/stable/30029508>.
- Giles, H., Mulac, A., Bradac, J. J., & Johnson, P. (1987). Speech Accommodation Theory: The First Decade and Beyond. *Annals of the International Communication Association*, 10(1), 13–48. <https://doi.org/10.1080/23808985.1987.11678638>.
- Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In *Contexts of accommodation: Developments in applied sociolinguistics* (s. 1–68). Editions de la Maison des Sciences de l’Homme. <https://doi.org/10.1017/CBO9780511663673.001>.
- Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1166–1183. <https://doi.org/10.1037/0278-7393.22.5.1166>.
- Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251–279. <https://doi.org/10.1037/0033-295X.105.2.251>.
- Goldinger, S. D., & Azuma, T. (2004). Episodic memory reflected in printed word naming. *Psychonomic Bulletin & Review*, 11(4), 716–722. <https://doi.org/10.3758/BF03196625>.
- Honeybone, P., & Watson, K. (2013). Saliency and the sociolinguistics of Scouse spelling: Exploring the phonology of the Contemporary Humorous Localised Dialect Literature of Liverpool. *English World-Wide. A Journal of Varieties of English*, 34(3), 305–340. <https://doi.org/10.1075/eww.34.3.03hon>.
- Kim, D., & Clayards, M. (2019). Individual differences in the link between perception and production and the mechanisms of phonetic imitation. *Language, Cognition and Neuroscience*, 34(6), 769–786. <https://doi.org/10.1080/23273798.2019.1582787>.
- Kim, M., Horton, W. S., & Bradlow, A. R. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, 2(1). <https://doi.org/10.1515/labphon.2011.004>.
- Machač, P., & Skarnitzl, R. (2009). *Principles of phonetic segmentation*. Epocha.
- Mitterer, H., & Ernestus, M. (2008). The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition*, 109(1), 168–173. <https://doi.org/10.1016/j.cognition.2008.08.002>.

- Namy, L. L., Nygaard, L. C., & Sauerteig, D. (2002). Gender Differences in Vocal Accommodation: The Role of Perception. *Journal of Language and Social Psychology, 21*(4), 422–432. <https://doi.org/10.1177/026192702237958>.
- Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics, 39*(2), 132–142. <https://doi.org/10.1016/j.wocn.2010.12.007>.
- Nielsen, K. (2014). Phonetic Imitation by Young Children and Its Developmental Changes. *Journal of Speech, Language, and Hearing Research, 57*(6), 2065–2075. https://doi.org/10.1044/2014_JSLHR-S-13-0093.
- Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America, 119*(4), 2382–2393. <https://doi.org/10.1121/1.2178720>.
- Pardo, J. S., Jay, I. C., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics, 72*(8), 2254–2264. <https://doi.org/10.3758/BF03196699>.
- Pardo, J. S., Jay, I. C., Hoshino, R., Hasbun, S. M., Sowemimo-Coker, C., & Krauss, R. M. (2013a). Influence of role-switching on phonetic convergence in conversation. *Discourse Processes, 50*(4), 276–300. <https://doi.org/10.1080/0163853X.2013.778168>.
- Pardo, J. S., Jordan, K., Mallari, R., Scanlon, C., & Lewandowski, E. (2013b). Phonetic convergence in shadowed speech: The relation between acoustic and perceptual measures. *Journal of Memory and Language, 69*(3), 183–195. <https://doi.org/10.1016/j.jml.2013.06.002>.
- Pardo, J. S., Urmanche, A., Wilman, S., & Wiener, J. (2017). Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics, 79*(2), 637–659. <https://doi.org/10.3758/s13414-016-1226-0>.
- Paquette-Smith, M., Schertz, J., & Johnson, E. K. (2022). Comparing Phonetic Convergence in Children and Adults. *Language and speech, 65*(1), 240–260. <https://doi.org/10.1177/00238309211013864>.
- Podlipský, V.J., Šimáčková, Š. (2015). Phonetic imitation is not conditioned by preservation of phonological contrast but by perceptual salience. *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK: University of Glasgow. Paper number 399. <http://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0399.pdf>.
- Posner, M. I. (1964). Information reduction in the analysis of sequential tasks. *Psychological Review, 71*(6), 491–504. <https://doi.org/10.1037/h0041120>.

- Rácz, P. (2013). *Saliency in Sociolinguistics: A Quantitative Approach*. Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110305395>.
- R Core Team. (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.r-project.org>
- Schertz, J., & Johnson, E. K. (2022). Voice Onset Time Imitation in Teens Versus Adults. *Journal of Speech, Language, and Hearing Research*, 65(5), 1839–1850. https://doi.org/10.1044/2022_JSLHR-21-00460.
- Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Schachar, M. S., Højsgaard, S., Fox, J., Lawrence, M. A., Mertens, U., Love, J., Lenth, R., Haubo, R., & Christensen, B. (2022). Afex: Analysis of Factorial Experiments. R package version 1.1-1. <https://CRAN.R-project.org/package=afex>.
- Thakerar, J. N., Giles, H., & Cheshire, J. (1982). Psychological and linguistic parameters of speech accommodation theory. *Advances in the social psychology of language*, 205, 205-255.
- Tobin, S. J. (2022). Effects of native language and habituation in phonetic accommodation. *Journal of Phonetics*, 93, 101148. <https://doi.org/10.1016/j.wocn.2022.101148>.
- Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag.

9 Annotation

Author: Anna Kovaříková

Field of Study: English Philology / General Linguistics and Communication Theory

Faculty and Department: Faculty of Arts, Department of English and American Studies

Supervisor: Mgr. Václav Jonáš Podlipský, Ph.D.

Number of characters: 68 297

Keywords: phonetic imitation, phonetic convergence, vowel duration, factors of phonetic imitation, phonological contrast

Description: This thesis studies how speakers imitate long vowel duration when stimuli they are exposed to follow short carrier phrases with a fixed speaking tempo and when the manipulation of the duration of these vowels threatens phonological contrast between the long vowel and the short vowel. It reviews literature relevant to potential influences of phonetic imitation. The experiment reported in this thesis specifically aims to replicate the experiment of Podlipský and Šimáčková (2015) which tested whether speakers imitate the extended duration of the Czech long vowel /u:/ when exposed to the extended duration of this vowel and whether they imitate the reduced duration of /u:/ when exposed to the reduced duration of /u:/. The results of the experiment suggest that contrast preservation plays a role in phonetic imitation as the participants imitated the extended /u:/, but the reduced condition was not imitated.

10 Anotace v češtině

Autor: Anna Kovaříková

Studijní obor: Anglická filologie / Obecná lingvistika a teorie komunikace

Fakulta a katedra: Filozofická fakulta, Katedra anglistiky a amerikanistiky

Vedoucí práce: Mgr. Václav Jonáš Podlipský, Ph.D.

Počet znaků: 68 297

Klíčová slova: fonetická imitace, fonetická konvergence, délka samohlásek, faktory fonetické imitace, fonologický kontrast

Charakteristika: Tato práce se zabývá tím, jak mluvčí imitují délku dlouhých samohlásek, když stimuly, kterým jsou vystaveni, následují krátké úvodní fráze se stálým tempem řeči a když manipulace délky těchto samohlásek ohrožuje fonologický kontrast mezi dlouhými a krátkými samohláskami. Práce představuje literaturu, která se zabývá potenciálními vlivy fonetické imitace. Experiment popsán v této práci konkrétně replikuje výzkum Podlipského a Šimáčkové (2015), který zkoumal, zda mluvčí imitují prodloužení délky dlouhé české samohlásky /u:/, když jsou vystaveni prodloužené délce této samohlásky, a zda imitují zkrácení /u:/, když jsou mu vystaveni. Výsledky experimentu naznačují, že snaha zachovat fonologický kontrast hraje roli ve fonetické imitaci, jelikož účastníci experimentu imitovali prodloužení /u:/, ale zkrácení /u:/ nebylo imitováno.