Filozofická fakulta Univerzity Palackého

Katedra anglistiky a amerikanistiky

Percepce hláskových kontrastů cizího jazyka dospělými posluchači

Návrh výzkumu percepce českých posluchačů anglického /w/ - / ɹ/

Perception of non-native sound contrasts by adult listeners

A research proposal on Czech perception of English /w/ - / ɹ/

(bakalářská práce)

Autor: Jana Hlušková (Anglická filologie – Německá filologie)
Vedoucí práce: Václav Jonáš Podlipský, Ph.D.
Olomouc 2012

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

V Olomouci dne 21. srpna 2012 .....

I am very thankful to Václav Jonáš Podlipský, Ph.D. for his willingness and patience as well as his helpful guidance and assistance with my thesis.

I am also very thankful to my schoolmates from the Department of English Philology for their useful information and moral support.

Contents

Abstract

During the first-language acquisition, humans learn to ignore irrelevant differences between different-sounding tokens of the same sound category of the ambient language. At the same time they learn to perceive with acuity linguistically relevant differences. Such higher-order modifications of phonetic perception do not (need to) occur for sound contrasts not used in the native language. The perception of a non-native contrast was discussed in this research paper. British English phonology contains the phonemes /w/ - voiced labio-velar approximant and /ɹ/ - voiced alveolar approximant. Czech phonology does not contain such a contrast. A research on adult monolingual Czech listeners' perception of the novel (English) /w/ - /ɹ/ contrast was proposed in this thesis.

1. Introduction

In this thesis I explore perception of a novel (English) sound contrast by monolingual adult Czech listeners. The aim is to find out whether monolingual adult Czech listeners can perceive the difference between the non-native phonemic /w/ - /ɹ/ contrast or they confuse one for another. According to the results of the proposed experiment we can make predictions about how the phonological system of the first acquired language affects the acquisition of the second-language sound categories.

Previous researchers (e.g. Kuhl, 2007; Gervain, Werker, 2008) suggest that during the first-language acquisition, people learn to ignore irrelevant differences between different-sounding tokens of the same sound category of the ambient language. At the same time they learn to perceive with acuity linguistically relevant differences in that language. Such higher-order modifications of phonetic perception do not occur for sound contrasts not used in the native language. Since both /w/ and /ɹ/ are foreign to a Czech ear, slight differences within each English category should be relatively easily perceptible for a monolingual Czech listener. At the same time no clearly defined boundary between /w/ and /ɹ/ should be found for monolingual Czech perceivers.

In this paper I will review literature focusing on sound-pattern learning. In section 1.1 I will describe the first language acquisition discussing also the phenomenon why infants master a language in general much easier than adults. As well, I will shortly outline the topic of a second-language acquisition as it is closely related to the previous two themes I just mentioned. In section 1.2 I will define cross-language perception basing this theory on the *Perceptual Assimilation Model* (PAM) designed by Best (1995) in section 1.2.1. Section 1.3 will deal with categorical perception. Next, in part 2. I will review the latest research on cross-language speech perception made by Best, Bohn (2011) and then in part 3. I will base my hypothesis and research questions on the theory I discussed in the previously mentioned parts and I will follow with proposing

my own research on cross-language speech perception in part 3.1. Finally, I will end my thesis with a discussion in section 4.

1.1. First language acquisition

As Gervain, Werker (2008) put it, speech perception and language acquisition have not been studied simultaneously from the beginning. They were not even considered to have so much in common. Nowadays we already know that infants' speech perception is helpful by acquiring abstract structural properties of language (Gervain, Werker, 2008). Perceiving speech and learning language are two linked highly sophisticated processes which infants appear to manage much easier than adults. "Despite their naiveté about the world in general, children can make and hear contrasts among dozens of speech sounds, they have learned thousands of words without having heard a single definition, and they are able to build and understand sentences of impressive complexity." (O'Grady, 2005: 2). Previous researches stated several reasons for it.

According to Kuhl (2004: 831) for speech perception infants require necessarily only basic perceptual capabilities which are also present in some animals, not only people. After that, when they are exposed to human speech, they learn very quickly in a way which I will describe and explain in the next paragraph. In order for children to learn language it is then necessary to be in contact with the language. It means that a social environment full of human interaction is required. A great lack of social interaction has an impact on future abilities of learning language (Kuhl, 2004: 831). In my personal opinion if an infant perceive only a very restricted number of stimuli, s/he will not be able to speak or understand the others. On the contrary infants who have a lot of people around them, thus a lot of human interaction, are usually more advanced in speaking their mother-tongue.

The most probable way in which infants learn language is called 'Statistical Learning' (Kuhl, 2007). This approach is based on discovering the patterns of a certain language and noticing the frequency of occurrence of the

units following each other. “Infants then rapidly learn from exposure to language, in ways that are unique to humans, combining pattern detection and computational abilities (often called statistical learning), with special social skills.” (Kuhl, 2007: 71). At first infants acquire the phonetic properties of particular language which supports further language development such as the word acquisition (Clark, 2009).

However, because of the lack of experience infants are also constrained in speech perception and language learning. “Children’s talent for language is strangely limited – they’re good at learning language, but not so good at knowing what to say and what not to say.” (O’Grady, 2005: 1). They are also not able to detect all of the patterns of language when they are exposed to speech and they are not capable of revealing every single physical difference in speech sounds. But although they cannot detect all of the physical differences and linguistic patterns, their selection of them is most important for language acquisition (Kuhl, 2004). I would call it “selecting”. It is as if the infants knew what to select in the first place and what they can shift for later. Thus before humans can even walk their brain can already work in an amazing way regarding the language perception.

Another factor which could explain why infants learn language with much less difficulty than adults is something called the “sensitive” or “critical” period (Tees, Werker, 1984) which is tightly connected with a so-called “neural commitment” (Kuhl, 2004). Already according to the name we can say that a “sensitive” period might be a state when infants are most alert to the inputs they perceive. And indeed it is so – infants are very sensitive to phonetic differences within and between all languages. By “neural commitment” it is meant that the neural cells of young infants gradually commit to language patterns which the infants have been most exposed to since birth (Kuhl, 2007).

Accordingly, the patterns of a certain language which the infant learns at first are decisive as these are the patterns s/he is going to concentrate on and develop. “This focussing in by infants towards the end of their first year on the phonetic categories of the specific language they are exposed to appears to

mark the first stages in the organization of a functional phonology, a sound system tied to a particular language." (Clark, 2009: 61). Later learning of patterns of a different language seems to depend on whether these patterns are compatible with the 'first learned' to which the infant has become "neurally committed" or not. This is a theory by Kuhl (2004) that explains the fact of not being able to learn the second language as properly as the first language by incompatibility of some of the second-language patterns to the first-language patterns. Let's look at the process of a first-language phonology acquisition in the following paragraph.

When acquiring the phonology of a language, infants at first learn the phones then they place them to identical or different phonological categories (Jusczyk, 1992). Distinct phones of the same phoneme are called allophones. These belong to the same phonological category. They just have different phonetic realizations. However, Eimas *et al.* (1971) proved that initially infants pay more attention to prosodic properties than to phonetic features of individual segments. Nazzi, Ramus (2003) support this idea in their paper on perceiving and acquiring the linguistic rhythm by infants. "However, infants' growing knowledge of their native language rhythmic properties seems to be the key to their ability to discriminate languages from within the native language rhythmic class by 5 months of age." (Nazzi, Ramus, 2003: 241).

An interesting point was made by researchers studying the relationship between early phonetic learning and later linguistic skills (Kuhl, 2007). Kuhl *et al.* (2005) came to the conclusion that the earlier infants are attuned to the native-language phones the more developed their linguistic abilities are at a later stage. And conversely, the longer they perform well in non-native sound distinction, the slower they evolve in their native language. This, I find very interesting as one would probably think that the longer humans are able to distinguish the native as well as the foreign sound contrast the better for the language skills development.

Although infants' language learning is an admirable process, we also have to take into account that while adults perceive the 'normal speed' speech,

infants are talked to with a reduced speech tempo and clearer articulation. This has been termed ‘motherese’ (Kuhl, 2007). People make these adjustments of their speech intuitively when addressing small children. Consequently it is easier for them to distinguish the differences between speech sounds (Kuhl, 2007). “It turns out that child-directed speech is often singularly well tailored to its addressees, highly grammatical in form, and virtually free of errors.” (Clark, 2009: 23).

An important thing to know is also the way people perceive the linguistic input. “Categorical perception is the tendency for adult listeners of a particular language to classify the sounds used in their languages as one phoneme or another, showing no sensitivity to intermediate sounds.” (Kuhl, 2004: 833). Infants are also alert to between-category sound changes with the difference, and at the same time the advantage, that at a very early age they are able to distinguish between speech sounds of all languages. As they approach the first year of age, they lose this universal ability and improve only the native language perception skills. They stop being the ‘perceptual citizens of the world’ and become ‘language specific perceivers’.

Now I will present a short list of infants’ speech development at particular age adapted from Jusczyk (1992):

- During the first month infants already perceive the speech sounds categorically.
- From one to four months of age they can reveal variation in intonation patterns and they are capable of switching attention in distinguishing some speech contrasts.
- From four to six months of age they reveal prosodic features in native and non-native language and react more to slower better articulated speech.
- From six to eight months of age they are able to differentiate the non-native from the native words according to prosody.

- From eight to ten months infants' capability of distinguishing a non-native phonetic contrast decreases.
- From ten to twelve months they have already shaped perceptual categories for their native language.

Here, I would shortly like to devote this paragraph to the second language acquisition which is related to the first language acquisition and the cross-language perception I am going to discuss in the next part. As we already said during the first year humans shape the perceptual categories for their native language. When humans start learning the second language, they are hindered by the system of the first language. According to Hakuta *et al.* (2003) with increasing age, the ability to learn a second language decreases. According to Hojen (2003) this decrease starts already at the age of 3-4. Hakuta (2003) also adds that by older people the ability to learn language also depends on their previous experience and education skills. This is, I think, clear because the more the head of a person is challenged and trained the more it is prepared for new information, as e.g. new language.

1.2. Cross-language speech perception

From what was said in the previous section, we could conclude that people are born with universal auditory receptiveness which adapts gradually to the sounds and patterns of the language which the person is primarily exposed to. According to Hojen (2003) humans start losing the universal auditory receptiveness when they are around 10 months old, depending on how different the foreign contrast is from their native dialect. It seems that this perceptual specialization driven by language experience is the reason why adults often have troubles with distinguishing non-native sound contrasts which do not have a phonemic value in their language. Previous hypotheses (Kuhl, 2007) claimed that this difficulty occurs because the neural sensors of a person already committed to different patterns s/he was exposed to since birth. However, the results of later research (Tees, Werker, 1984) indicated that the

reason for adults not to be able to discriminate non-native contrastive sounds is their specific phonemic processing approach applied on their mother tongue.

There had to be large manipulations in the length of the interval between the speech signals or particular testing procedure used so that the examined listeners could change their native phonemic processing strategy and be able to perceive the within-category sounds. Then they would be able to discriminate the non-native sound contrast using a new linguistic processing approach.

Regarding the universal auditory sensitivity, the sensorineural reactiveness is not lost, as adults proved the ability of discrimination of the non-native sound contrast in special trial circumstances (Tees, Werker, 1984). Tees, Werker (1984) further suggests that they are just not able to benefit from this ability when examined in classical demanding conditions used for speech perception. Here we can see that the tempo of a common native speech plays also quite a big role in making the perception more difficult.

However, there is another view Hojen (2003) discusses in his thesis. It concerns the phonetic variations of a phoneme (allophones) versus the phonemic distinctions between languages which the second language listeners are supposed to perceive. This view makes us clearer the big difference between the phonemes and its variations. "The difficulties that L2 learners have perceiving L2 sounds differ across different allophones of a phoneme. This suggests that the unit of analysis in the perception of speech sounds is not the phoneme. Rather, listeners are sensitive to the phonetic details of the L2 sounds which differ according to phonetic context." (Hojen, 2003: 14). I consider this theory very interesting as it suggests another point of view to the phenomenon of a non-native contrast perception. This point of view looks deeper into the nature of phonemes and proposes that we have to consider also their variations. This theory I find quite detailed and sophisticated.

As we already know from the section of the first-language acquisition, humans learn their native languages by attuning their perceptual system to the relevant sound differences and patterns and by learning to systematically ignore the distinctions that are not relevant in that particular linguistic system. After the perceptual system has been tuned and the human has picked up appropriate distinctive information at the lower-order level (phonetic segments), s/he starts eliciting information at higher-order level such as the syntactic structure. “Spoken language provides an excellent example of the sort of complex organization in which higher-order invariants, such as those that specify syntactic structure, may not be detectable until the perceiver has learned to pick up certain distinctive information at lower levels.” (Best, 1995: 184). Thus, just to explain Best’s (1995) words, establishing higher-order units of the native language increases the inability of extracting lower-order properties of the unknown non-native categories not used contrastively in the mother tongue. When higher-order units are recognized listeners need less lower-order information from the perceived stimuli and they can ignore the remaining phonetic input.

What is it that makes sound patterns of different languages so different? Best (1995) claims that languages vary most markedly in the way they group simple gestures into more complex gestural constellations. Another crucial factor is the typical phasing relationships between these gestures in different languages. Hojen (2003) devoted a part of his thesis to the fact how foreign accent is detectable in second language learners. He states that even L2 learners at the age of 3-4 may already have a detectable foreign accent. The reasons of the detectability of the accent are the segmental mistakes as well as a different prosody. “It has often been reported that the degree of foreign accent correlates with, among other dimensions of L2 speech, the amount and gravity of segmental errors.” (Hojen, 2003: 7).

Although we pointed out the cross-language differences in sound categories and patterns we should probably make clear that this does not mean that languages have nothing in common phonetically. In fact because of the

particular shape, biological structure and constraints of the human vocal tract, languages contain many identical or similar sounds as they have equal possibilities for sound productions (Best, 1995). This factor should not be forgotten as so far I was just talking about the differences between the native and non-native sounds but we have to keep in mind that there are also features that they have in common. I could adopt the human vocal tract to a vase and the sounds of various languages to different flowers (different kinds and different colours). It is as if we were changing the flowers in the same vase.

Let's look at the model proposed by Best (1995) which should serve for understanding how the perception of a non-native contrast may work.

1.2.1 Perceptual Assimilation Model (PAM) designed by Best (1995)

"Non-native segments are those whose gestural elements or intergestural phasing do not match precisely any native constellations." (Best, 1995: 193).

The basis of the perceptual assimilation model of cross-language speech perception is that listeners incline to perceive non-native sounds in accordance with the similarity to or difference from the native sounds (e.g. according to the spatial closeness of constriction locations and active articulators and resemblances in degree of constriction and gestural phasing). It means that the perceiver is assumed to reveal gestural resemblances to native phonemes. Moreover, s/he will most likely also detect differences from the native phonemes, particularly if they are very big.

In such a case listeners may not even find a native category to which they could assimilate the non-native phone, they only recognize it as a speech-like sound. Furthermore, in the most extreme case these non-native phonetic gestures do not even have to be perceived as speech-like sounds at all. Now three levels of assimilation of non-native *segments* suggested by the perceptual assimilation model will be presented:

1) *Assimilation to a native category:* the non-native segment resembles sounds from particular native category. It will be perceived as a native-like phone.

2) *Assimilation as uncategorized speech sound:* has a speech-like resemblance but is not similar enough to any native category. It will be perceived as uncategorized speech sound.

3) *Not assimilated as speech sound (non-speech sound):* does not even resemble speech. Therefore perceived as non-speech sound.

However, more important for this paper is how the non-native *contrasts* will be perceived. This depends, naturally, on how the two members of the contrasting pair are perceived individually. PAM distinguishes six types of a non-native contrast assimilation listed below. Thanks to these types, discrimination of each contrast can be predicted:

a) *Two-Category Assimilation* – each of the non-native phones is assimilated to a distinct native category. Expectations for discrimination are very high.

b) *Category-Goodness Difference* – the two non-native segments are both assimilated to the same native category but they differ in the degree of similarity. One of them is more similar to the native sound than the other. Expectations for discrimination are moderate to high.

c) *Single-Category Assimilation* – both non-native sounds are assimilated to an identical native category and the degree of similarity to the native segment is equal for both. Expectations for discrimination are low.

d) *Both Uncategorizable* – each of the non-native segments is realized as speech-like sound. However, none of them is assimilated to a native category. Discrimination should depend on the closeness of these two sounds as well as on the closeness

to native categories. Thus discrimination can probably prove poor to very good.

e) *Uncategorized versus Categorized* – one of the non-native segments is assimilated to a native category and the other is not. However, it is at least recognized as speech-like element. Expectations for discrimination are high.

f) *Nonassimilable* – both of the non-native segments are perceived as non-speech sounds. Expectations for discrimination are high to very high.

My hypothesis is going to be based on PAM as I consider it a very progressive model of human perception in the field of perceptual phonology.

At this point I would like to look a little bit into the relation between the perception and production of a foreign contrast. In the research by Bradlow *et al.* (1997) on the relation between the perception and production of a foreign contrast, the researchers found out that a perceptual training supports production. Thus, after a perceptual training on the English /r/ - /l/ sound contrast the Japanese speakers' production of /r/ and /l/ was much clearer than before the training. However, I would like to add from my own experience in a foreign environment with a foreign language that although perception may support production but only a proper training 'codes' the way of the correct producing in your brain.

1.3. Categorical perception

"The term 'categorical perception' refers to the fact that labeling performance predicts discrimination: adults better discriminate pairs of stimuli of equal sized differences if the two stimuli cross their labeling boundary than they do if the two stimuli are from within a single phoneme category." (Gervain, Werker, 2008: 1157). This definition indicates the way humans perceive speech.

Categorical perception, which has already been outlined above, is a very powerful phenomenon occurring during the speech perception in infants as well as adults. In many experiments people discriminated better equal physical differences across the phonemic boundary than within the same phonemic category. Previous findings which came from an experiment on speech perception in 1 and 4 month old infants suggest that categorical perception is probably a part of the biological composition of a human being that starts working at a very early stage of life (Eimas *et al*., 1971).

In previous research on speech continua (e.g. Eimas *et al*., 1971) beginning from one consonant phoneme heading in equal acoustic steps to another (e.g. /p/.../b/) researchers found out that these stimuli are perceived rather discontinuously. As Strange (1999) states it, sounds are, of course, determined acoustically. For example approximants have more gradual formant transitions than plosives. Experiments on speech perception have often been based on manipulating the acoustic properties of speech sounds in order to create a continuum of equal steps heading from one phonetic category to another (Strange, 1999). Some adjacent stimuli were perceived as equal while other adjacent stimuli differing in the same magnitude from one another were perceived as different phonemes (Strange, 1999).

According to these findings we can say that consonants are perceived categorically. As it was already said in section 1.1 this is because of the boundary that speakers have created for the phonemes of their native language. When perceiving speech it is very helpful that people can handle acoustic phonetic differences within the same phonemic category and label them as identical. Thanks to this capability we are able to identify words as the same quickly despite distinct pronunciations (Jusczyk, 1992).

Experiments on categorical perception include two kinds of tests following each other. Strange (1999) describes them in his paper. The first one is the *identification test* based on labeling each of the sounds of the non-native sound contrast. Shortly, the participants are played several stimuli of the non-native contrast in random order more times and they are supposed to say which

of the sounds of the non-native pair it was. The second one is the *discrimination test* which focuses on detecting the phonetic differences between two members of a continuum that are always equally distant. Predicting that listeners will perceive this contrast categorically, according to the results of the identification test we should be able to foretell the results of the discrimination test, in other words the discriminability. Typical results of *identification tasks* and *discrimination tasks* is that the listeners perceive the steps of the continuum between the native sounds discontinuously and you can already predict a result of a discrimination test from the result of the identification test. That means according to an *id task* where usually perceivers of the native contrast divide the continuum into several categories depending on how many phones they are supposed to label, a precise discrimination for cross-category pairs and quite a poor discrimination for within-category pairs is predicted.

I would say, basing my thoughts on the things we already discussed above, that these tests have such results because of the fact that in the infant period humans create phonemic categories for their mother-tongue and therefore they stop recognizing the within-category differences of this language.

"Categorical Perception refers to the fact that perception of an acoustic cue that varies along a continuum of equal steps is discontinuous and that the discontinuities correspond to the boundaries between phoneme categories." (Strange, 1999: 172). Categorical perception is proved by associated identification and discrimination tests. Strange (1999) gives four characteristics:

1) Identification of a speech continuum indicates a steep slope. This is due to the extreme change near the boundary of a phonetic category.

2) Precise discrimination appears just at the category boundary.

3) Within-category discrimination is very inaccurate. Mostly based on guessing.

4) Identification can foretell discrimination at every point on the continuum.

If all of these criteria are valid, categorical perception has its full extent. It happens very often, however, that only the first two points of this scale are met. In such a case we are speaking about the 'Phoneme Boundary Effect' (Strange, 1999).

2. Review of the latest research paper on perception of a non-native contrast

Here I will review the latest research made by Bohn, Best (2011), where the phonological as well as the phonetic impact of the native language on perception of several non-native consonant contrasts was tested. German and Danish native speakers perceived American English consonants /r, l, w, j/. Both of these languages have all of these consonants except for /w/. Therefore they lack the phonological contrasts /w/-/j/ and /w/-/r/. On the contrary, they distinguish /r/-/l/ as phonological contrast. Regarding the phonetic factor German and Danish employs light (alveolar) /l/ while American English uses dark (velarized) /ɬ/. /j/ is realized in all of these three languages identically. Furthermore, English /r, w/ resembles German and Danish /r, v/ but Danish realisations of /r, v/ are more similar to the English /r, w/ realisations.

The results showed that the phonetic factor had a greater influence on perception of these sounds than the phonological factor. Danish listeners performed on /w-/r/ and /r/-/l/ almost as good as the native English speakers and they discriminated /w/-/j/ even better than English speakers while German listeners identified all the contrasts highly categorically. However, their discrimination of /w/-/r/ and /r/-/l/ was significantly worse than that of Danish and English listeners. Their discrimination of /w/-/j/ was better than by English speakers but worse than by Danish speakers. Best, Bohn (2011) stated, therefore, that the phonetic resemblances have a greater impact on cross-language speech perception than the phonological equivalents.

Later on, in section 4. I will make my own research proposal digging further into the cross-language perception and categorical perception finding the way to discover the influence of the native (Czech) language on the non-native English /w/- /ɹ/ perception.

3. Hypothesis and research questions

My hypothesis and research questions are based on the information about the first language acquisition, cross-language perception and categorical perception which I was discussing in the introductory part.

Will the Czech adults be able to identify the non-native /w/ - /ɹ/ contrast as two distinct speech sound categories? Will they be able to discriminate different members of a /w/ - /ɹ/ speech continuum? How will (according to PAM) the Czech listeners assimilate the non-native sound contrast? Will they perceive the contrast categorically?

British-English phonology contains the phonemes /w/ - voiced labio-velar approximant and /ɹ/ - voiced alveolar approximant. As we can see these phonemes vary only in the place of articulation. Czech phonology does not have the phonemes /w/ and /ɹ/ but it has /r/ - voiced alveolar trill and /v/ - voiced labiodental fricative which are the closest phonemes to the foreign contrast. Thus Czech language does not distinguish the English /w/ - /ɹ/ contrast. It means that Czech /r/ and English /ɹ/ have the same voicing and place of articulation while Czech /v/ and English /w/ only have the same voicing and the lips are present in producing both of these sounds but in connection with a different second articulator (/v/ - the teeth; /w/ - the velum).

According to what was said about the *first language acquisition* in section 1.1., by the time humans are one year old they have already shaped the perceptual categories for their native language. I, therefore, presume that Czech adults will not be able to identify the non-native /w/ - /ɹ/ contrast as two distinct speech sound categories since they don't have them in their L1. Their neural cells already committed to the native language patterns which do not

include this phonemic distinction (/w/ - /ɹ/). However, there is a chance that to them this contrast will appear similar to the native /v/ - /r/ contrast. In this case they should be able to identify the non-native sound contrast as two distinct speech sound categories. On the contrary, they should be able to discriminate different members of the /w/ - /ɹ/ continuum quite well because in their L1 development they never had to learn to classify a number of /w/ sounding tokens as members of one category (ignoring the within-category differences) and a number of /r/ sounding tokens as members of another category. They have no such categories (unless they use the Czech /v/ and /r/ categories). Thus even the differences which would be imperceptible ("ignored") for English perceivers as these are the within-category differences (the English category of /w/ or of /ɹ/) should be noticeable for the monolingual Czechs.

In relation with the topic of *cross-language perception* we were talking about PAM (Perceptual Assimilation Model) designed by Best (1995). PAM introduces us the levels of assimilation of the non-native segments to the native ones. As we already said earlier Czech listeners could assimilate English /w/ to their native /v/ and the non-native /ɹ/ to Czech /r/. It depends on the degree of resemblance of these non-natives with the native segments for Czech listeners. If they are similar enough for them, the non-native sounds will be perceived as the native-like sounds. On the contrary, if they are not similar enough for them, they will perceive them as uncategorized speech sounds. That means that these non-native phones will be realized as speech sounds but not as the native category sounds /v/ and /r/. According to the fact that the English phoneme /ɹ/- voiced alveolar approximant and the Czech phoneme /r/ - voiced alveolar trill have more in common than English /w/ - voiced labio-velar approximant and Czech /v/ - voiced labiodental fricative the assimilation degree by /ɹ/ and /r/ is higher. However, what is more important for our research is how the whole sound contrast will be perceived.

PAM suggests six cases for the assimilation of the sound contrasts, which, of course, primarily depend on how the sound members of each contrast are perceived individually. I would say, for the non-native /w/ - /ɹ/ contrast to Czech perceivers, three situations could be relevant: *Two-Category*

Assimilation - /w/ and /ɹ/ will be assimilated to a distinct native category (e.g. /w/ to /v/ and /ɹ/ to /r/). The discrimination, in such case, is expected to be very high; *Both Uncategorizable* - neither /w/ nor /ɹ/ will be assimilated to a native sound category. However, it should be mentioned that both are realized as speech-like sounds. The predictions for discrimination are poor to very good depending on the closeness of the non-native phones to each other as well as on the closeness to the native categories; *Uncategorized versus Categorized* – e.g. when /ɹ/ will be assimilated to the native /r/ but /w/ will only be realized as a speech-like sound but not assimilated to /v/ or vice versa (considering the first case more probable according to the degree of similarity of the phonemic descriptions I have already mentioned above). The discrimination is predicted to be high.

The basic description of the phenomenon 'categorical perception' we were talking about in the introduction is that phones across the phonemic boundary are perceived much easier than the segments within category. That means, as we already discussed in section 1.3., that sounds are perceived categorically. This applies for the native sound contrasts, though. At birth infants are able to perceive also the within-category differences of all languages. Then, during the exposure to the native language they shape the phonemic categories for the native language. After that they ignore the irrelevant differences of the within-category phones and perceive only relevant differences of the between-category phones of the ambient language.

However, /w/ and /ɹ/ are unknown to a Czech ear as it is a non-native contrast, so subtle differences within each English category should be relatively easily perceptible for a monolingual Czech perceiver. Simultaneously, no precisely defined boundary between /w/ and /ɹ/ should be found for monolingual Czech listeners. Thus, according to the theory of *categorical perception,* I predict that the results of the identification and discrimination tests will be the contrary of categorical perception. Thus the sounds will probably not be perceived categorically. Therefore, the criteria of categorical perception defined by Strange (1999) will most likely be reversed: the identification of a speech continuum will not indicate a steep slope due to

the extreme change near the boundary of a phonetic category, as there is no such category. Precise discrimination will not appear just at the category boundary. Within-category discrimination will be accurate. According to the identification we cannot foretell the discrimination at every point on the continuum.

3.1. Research proposal

The goal of the experiment is to find out how the adult Czech listeners will perceive the non-native English /w/ - /ɹ/ contrast. The cross-language perception based on the theory of categorization is tested.

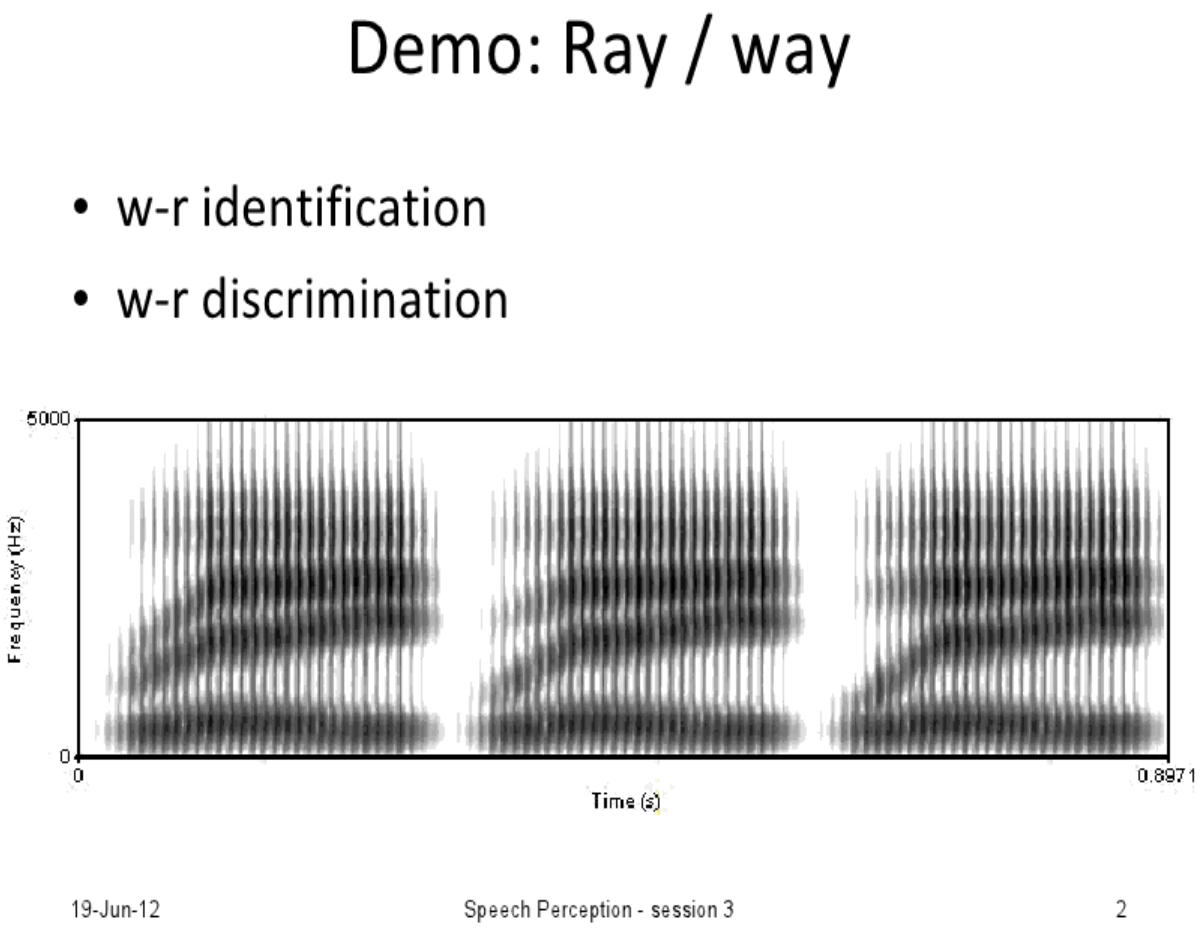
Method:

Subjects:

There should be at least one naive adult Czech listener tested in this experiment, no matter what sex and what age. Here I should probably say that the more participants there are the more precise the results will be. A condition for selecting the participant/s is that they have to be monolingual, that means having Czech parents, speak only Czech with them and having no knowledge of other language than Czech and especially, no knowledge of English language, that means never having learned English and never having been in an English-speaking country, is expected. All of the participants must fulfill these conditions completely. To evaluate and be able to compare the perception of a non-native sound contrast there should be a control group of native British listeners. Again no matter what sex and what age. These native speakers are also given the same test on perception of the /w/ - /ɹ/ sound contrast.

Picture 1: Speech perception testing

Changing the F3 formant frequency from /w/ to /ɹ/ creating a continuum of a particular number of steps.



Stimulus material:

A seven-step way – ray speech continuum (/w/ - /ɹ/ continuum in the same context) would be used. The speech is originally recorded from a native British English speaker and already synthesized in equal steps changing the F3 formant frequency. The *Open Source Software for Experiment Design and*

Control described by Gayvert, Hillenbrand (2004) would be used for the controlled listening experiment which includes the identification- and discrimination test.

"The software runs under Windows and is controlled by creating text files that allow the experimenter to specify key features of the experiment such as the stimuli that are to be presented, the randomization scheme, interstimulus and intertrial intervals, the format of the output file, and the layout of response alternatives on the screen." (Gayvert, Hillenbrand, 2004: 45).

Procedure:

Firstly, the Czech listener/s is/are given an *Identification test* described e.g by Strange (1999). S/he is supposed to label the 8 stimuli which are played several times in random order as /w/ or /ɹ/. Then s/he is given a *Discrimination test* (Strange, 1999) in which s/he is supposed to detect any difference between the by-standing steps of the continuum from /w/ to /ɹ/ differing in F3 by equal amount. To make it easier for the Czech perceiver s/he is given an *AB X Discrimination test*. After a pair of stimuli AB one of these stimuli is repeated as X. The subject's task is to say whether X is the same as A or B. The native (English) control group have the same tasks as the Czech listener/s. After the performance of the subjects a comparison of the results will be made.

4. Conclusion

According to the results of the identification and discrimination test taken by the Czech and English listener/s we will be able to see the difference between the perception of /w/ - /ɹ/ as a native contrast (by English perceivers) and the perception of /w/ - /ɹ/ as a non-native contrast (by Czech perceivers). We will find out whether the Czech adults are able to identify the non-native /w/ - /ɹ/ contrast as two distinct speech sound categories. Moreover, we will discover if they are able to discriminate different members of a /w/ - /ɹ/ speech continuum and how they assimilate the non-native sound contrast. Finally, we will see if the adult Czech listeners perceive the contrast categorically.

English speakers will probably perceive this contrast categorically as they are used to this sound contrast from birth. The perception of this sound contrast by Czech speakers will probably be the contrary of categorical perception.

Because of the fact that both /w/ and /ɹ/ are foreign to a Czech ear, subtle differences within each English category should be relatively easily perceptible for a monolingual Czech listener. Simultaneously, no clearly defined boundary between /w/ and /ɹ/ should be found for monolingual Czech perceivers.

In conclusion, I found it very interesting to get a lot of new information from the research articles on speech perception and language acquisition and discuss them. I am sorry, not having made myself the research I proposed but unfortunately I did not find a person who would meet all of my conditions. Despite this fact working on this thesis brought me a lot for my future studies.

English summary

The perception of a non-native sound contrast by adult monolingual Czech listeners was dealt with in this thesis. Concretely, the foreign (English) sound contrast /w/ - /ɹ/ was explored and discussed. In the theoretical part I mainly commented on and reviewed the literature from the three phenomena - first language acquisition, cross-language speech perception and categorical perception. I also touched the topic why infants master language easier than adults and I devoted a standalone section to the short outline of the second-language acquisition. As for the cross-language speech perception I concentrated on the Perceptual Assimilation Model (PAM) designed by Best (1995). Furthermore, I reviewed the latest research paper on the perception of a non-native sound contrast written by Bohn, Best (2011).

While acquiring the first language people learn to perceive only the linguistically relevant differences for the ambient language. They also learn to ignore the irrelevant differences between different-sounding tokens of the same category. This ability does not necessarily need to work with the non-native contrasts of the second language.

For the practical part I made my own research proposal on perception of the novel (English) sound contrast. British English phonology has the phonemes /w/ - voiced labio-velar approximant and /ɹ/ - voiced alveolar approximant. Czech phonology does not include such a contrast. The goal was to find out whether the monolingual adult Czech listeners can perceive the difference between /w/ and /ɹ/ that do not have a phonemic value in their language. The results of the proposed experiment would help us to be able to make predictions about how the phonological system of the first acquired language affects the acquisition of the second-language sound categories.

České shrnutí

Tato práce se zabývá vnímáním cizího zvukového kontrastu českým/i monolinguálními posluchači. Konkrétně se zaměřuji na zkoumání a projednávání zahraničního (anglického) zvukového kontrastu / w / - / ɹ /. V teoretické části zkoumám hlavně tři lingvistické jevy - akvizici prvního jazyka, mezijazykovou řečovou percepci a kategorickou percepci. Také se dotýkám tématu, proč děti zvládají jazyk jednodušší než dospělí. Samostatný odstavec pak ještě věnuji krátkému načrtnutí akvizice druhého jazyka. Pokud jde o medzijazykovou řečovou percepci, zaměřila jsem se na model perceptuální asimilace (PAM) který navrhla autorka Best (1995). Dále jsem zrekapitulovala článek nejnovějšího výzkumu na percepce cizího zvukového kontrastu, autormi kterého jsou Bohn, Best (2011).

Při akvizici prvního jazyka se lidé učí vnímat pouze jazykově významné rozdíly pro tento jazyk. Učí se také ignorovat irelevantní rozdíly mezi různě znějícími prvky ze stejné kategorie. Tato schopnost nemusí nutně fungovat s cizími kontrasty druhého jazyka.

Pro praktickou část jsem navrhla svůj vlastní výzkum zaměřený na percepci cizího (anglického) zvukového kontrastu. Fonologie britské angličtiny má fonému / w / - znělý labio-velární approximant a / ɹ / - znělý alveolární approximant. Ovšem česká fonologie takový kontrast neobsahuje. Cílem tooto výzkumu je zjistit, zda monolinguální dospělí čeští posluchači jsou schopní vnímat rozdíl mezi / w / a / ɹ /, které nemají fonologickou hodnotu v jejich jazyce. Výsledky navrženého experimentu by nám mohli pomoct předpovědět, jak fonologický systém prvního získaného jazyka ovlivňuje osvojování si zvukových kategorií druhého jazyka.

References

Best, C. T. (1995). A direct realist view of cross-language speech perception. In Strange, W. (Ed.), Speech perception and linguistic experience: Issues in Cross Language Research. Baltimore: York Press, 171-206.

Best, C. T., and Bohn, O. S. (2011). Native-language phonetic and phonological influence on perception of American English approximants by Danish and German listeners. Aarhus University: Elsevier Ltd. doi: 10. 1016/j.wocn.2011.08.002.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. Indiana University, Bloomington, Indiana: Acoustical Society of America. Accessed April, 1997.

Clark, V., Eve. (2009). First Language Acquisition. Stanford University: Cambridge University Press.

Eimas, P. D., Einar, S. R., Jusczyk, P., and Vigorito, J. (1971). Speech perception in infants. American Association for the Advancement of Science, 303-306. Accessed March, 2011.

Gayvert, R.T., and Hillenbrand, J.M. (2004). Open Source Software for Experiment Design and Control. American Speech – Language – Hearing Association, 45 - 60. doi: 10.1044 / 1092 – 4388(2005 / 005).

Gervain, Judit, and Janet F. Werker. (2008). How Infant Speech Perception Contributes to Language Acquisition. Language and Linguistics Compass, 1149-1170. doi: 10.1111/j.1749-818X.2008.00089.x

Hakuta, K., Bialystok, E., and Wiley, E. (2003). Critical Evidence: A Test of the Critical-Period Hypothesis for Second Language Acquisition. Stanford University and York University, Toronto, Ontario, Canada: American Psychological Society. Accessed January, 2003.

Hojen, D., Anders. (2003). Second-language speech perception and production in adult learners before and after short-term immersion. University of Aarhus, Department of English.

Jusczyk, W., Peter (1992). Developing phonological categories from the speech signal. In Ferguson, C. A., L. Menn, and C. Stoel-Gammon (eds.), Phonological Development: Models, Research, Implications. Timonium, MD: York Press.

Kuhl, K., Patricia. (2004). Early language acquisition: cracking the speech code. Nature Reviews Neuroscience 05: 831 - 843. Accessed November, 2004. doi: 10.1038/nrn11533.

Kuhl, P. K., Conboy, B. T., Padden, D., Nelson, T., and Pruitt, J. (2005). Early Speech Perception and Later Language Development: Implications for the ″Critical Period″. Psychology Press, 237 - 264. doi: 10. 1080/ 15475441.2005.9671948.

Kuhl, K., Patricia. (2007). Cracking the speech code: How infants learn language. PACS 4. Accessed February, 2007. doi: 10.1250/ast.28.71.

Nazzi, T., and Ramus, F. (2003). Perception and acquisition of linguistic rhythm by infants. Speech Communication 41 (2003) 233 – 243.

O' Grady, William. (2005). How Children Learn Language. University of Hawai: Cambridge University Press.

Strange, Winifred. (1999). Perception of Consonants: From Variance to Invariance. York Press, 167-182.

Werker, J. F., and Tees, R.C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. Acoustical Societyof America.

Anotace

Autor: Hlušková Jana
Katedra: Katedra anglistiky a amerikanistiky FF UPOL
Název česky: Percepce hláskových kontrastů cizího jazyka dospělými posluchači
Název anglicky: Perception of non-native sound contrasts by adult listeners
Vedoucí práce: Václav Jonáš Podlipský, Ph.D.
Počet stran: 31
Počet znaků: 50 937
Počet příloh: 0
Počet titulů použité literatury: 17
Klíčová slova v ČJ: percepce, kategorická percepce, mezijazyková řečová percepce, akvizice prvního jazyka, domácí / cizí zvukový kontrast
Klíčová slova v AJ: perception, categorical perception, cross-language speech perception, first language acquisition, native / non-native sound contrast

Anotace v ČJ: Inspirací pro tuto práci bylo zjistit spoustu nových informací o vnímání cizího zvukového kontrastu pro monolinguální dospělé české posluchače. Hlavním cílem této práce je podívat se hlouběji do tří jazykových jevů: akvizice prvního jazyka, mezijazykové řečové percepce a kategorické percepce. Je zde vytvořen také návrh výzkumu, který je postaven na těchto jevech. Účelem je zjistit, zda monolinguální dospělí čeští posluchači dokážou identifikovat a rozlišit anglický zvukový kontrast / w / - / ɹ /.

Anotace v AJ: The inspiration for this thesis was to find out a lot of new information about the perception of a non-native sound contrast for monolingual adult Czech listeners. The main goal of this paper is to look deeper into three linguistic phenomena: the first language acquisition, cross-language speech perception and categorical perception. There is also a research proposal based on these phenomena in this thesis. The purpose is to find out whether the monolingual adult Czech perceivers can identify and discriminate the English /w/ - /ɹ/ sound contrast.