“Never too late to spot a difference”: the development of L2 speech perception in adults

Dorota Lipińska

University of Silesia, Poland

Abstract

Learning correct pronunciation of a second/foreign language always represents a considerable challenge for language learners (e.g. Rojczyk, 2010a), especially for adults (e.g. Flege, 2007). There is an abundance of studies (e.g. Nowacka, 2010; Flege, 1991) showing that second language learners whose first language (L1) phonetic system has only one sound where L2 is characterized by noticeable richness of separate sound categories, encounter serious problems when they try to distinguish those new sounds and, moreover, they tend to apply their native vowels or consonants in L2 speech. It may be easily audible in the case of vowels and actually a lot of studies on L2 learners' production and perception of L1 and L2 vowels have been carried (e.g. Flege, 1992; Nowacka, 2010; Rojczyk, 2010a; Rojczyk, 2010b).

The aim of this study was to examine elementary learners' perception of 4 German vowels, namely: /ɪ/, /i:/, /ʏ/ and /y:/. They were organized as two sets of minimal pairs, namely /ɪ/ vs. /ʏ/ and /i:/ vs. /y:/. The aforementioned sounds were chosen for the study since /ʏ/ and /y:/ are considered to be very difficult vowels for Polish learners (e.g. Bęza, 2001). Twelve elementary, adult (29-52 years old) Polish learners of German agreed to participate in the study. The subjects had just began their A2-level language course, however they had been taught the basics of German pronunciation for a year, during their A1-level course. They were presented a printed list of word pairs and listened to the recorded words. Then they were asked to circle the right option in each pair. The results revealed that although all study participants were adults when they started learning German and they were still just elementary users of the language, they were already able to distinguish correctly a considerable number of words. It may suggest that proper pronunciation training during a FL/L2 course can provide language learners with measurable benefits.

1. Introduction

In recent years second/foreign language teaching has been focused on successful communication. It has included concentration on conversation skills and on reduction of the amount of formal instruction in other, “traditional” skills. However, in order to communicate successfully, a language learner has to, first of all, understand other speakers. Moreover, language users have to be understood correctly and their speech must be intelligible enough to convey the intended message (e.g. Littlewood, 1994). Although L2 learners still care a lot not only about basic communication, but also about grammatical (syntactic) norms and errors at all stages of proficiency, they frequently tend to forget that grammatical norm is not the only type of norm which ought to be taken into account if one wants to approximate (more or less) the native models, and they often disregard pragmatic, morphological, orthographic and phonetic norms (Sobkowiak, 2004). It is a common situation when L2 learners care less for proper pronunciation and usually pay more attention to comprehension skills and grammatical rules, especially when they have not been trained to discriminate major phonetic contrasts since the early stages of learning their L2 (Eddine, 2011).

2. Perception of L2 sounds

Learning a new sound system is always complicated (e.g. Rojczyk, 2010a). Language teachers have observed similar pronunciation difficulties which have been later confirmed in numerous studies. A lot of research (such as Nowacka, 2010; Flege, 1991) proved the hypothesis that L2 learners whose first language (L1) sound system is characterized by only one sound where L2 has two (or more) separate sound categories, encounter problems to distinguish new sound categories and, moreover, tend to apply their L1 sound to both new contexts. There are two models (Speech Learning Model – SLM, Flege, 1995, and Perceptual Assimilation Model – PAM, Best 1994) predicting that in the case when an L2 learner’s brain encounters two new sounds which are “similar” to the one, already existing L1 sound category, it will have difficulties in separating them.

The SLM predicts that phonetic similarities and dissimilarities of the learner’s L1 and their TL’s segments shall influence the degree of success in production and perception of non-native sounds as bilinguals are not able to fully separate their L1 and L2 phonetic subsystems (Flege et al., 1997). The sounds in L2 are divided into two kinds – “new” and “similar”. “New” sounds are those which are not associated by learners with any L1 sound. “Similar” sounds are those regarded as the same as certain L1 sounds (Brown, 2000). The phonetic similarity and dissimilarity are defined in terms of the acoustic and articulatory characteristics of the linguistically relevant speech sounds. The attainment of native-like production and perception of given second-language sounds is connected to the phonetic distance between L1 and L2 segments (Flege, 1995; Rojczyk, 2010a). Although the phonetic systems participating in perception tend to remain adaptive over the lifespan and reorganise in response to sounds in the L2 input, the process of “equivalence classification” makes it difficult or even impossible to establish new phonetic categories for sounds which are similar (Brown, 2000). Thus it is thought that L2 learners will be less successful in learning those L2 sounds which are regarded as being similar to L1 sounds (Flege, 1995; Rojczyk, 2010a).

The Perceptual Assimilation Model claims that the difficulties which L2 learners encounter while learning L2 speech sounds are determined by perceptual limitations. The PAM suggests that L2 listeners tend to classify sound contrasts in L2 into various categories, depending on the degree of similarity between their native and non-native sounds (Pilus, 2005). Similarity is understood here as the spatial proximity of constriction location and active articulators (Brown, 2000). Such classifications of L2 contrasts do determine how these

contrasts will be assimilated to learners' native categories. Generally speaking, the fact that L2 sounds are perceived as speech or non-speech sounds constitutes a basis for classification of L2 contrasts. L2 learners categorise L2 contrasts which are perceived as speech sounds differently than those perceived as non-speech sounds. L2 contrasts classified as speech sounds are then divided into two categories: native and non-native (new). It depends on whether they can assimilate to already existing native categories or not (Best, 1995; Pilus, 2005).

According to the aforementioned hypotheses and similar theories (such as Markedness Differential Hypothesis – Eckman, 1977), numerous studies have been conducted which aimed at examining L2 learners' successes and failures in production and perception of L1 and L2 vowels (e.g. Flege, 1992; Flege et al., 1997; Nowacka, 2010; Porzuczek, 2007; Rojczyk, 2009; Rojczyk, 2010a; Rojczyk, 2010b; Lipińska, 2013a; Lipińska, 2013b). Most of them have proved that it is difficult for L2 learners to separate “old” and “new” vowel categories, both in production and perception.

3. Age factor

Learners' age is one of the crucial factors enumerated while discussing the process of second language acquisition. Here one ought to mention the Critical Period Hypothesis. The hypothesis was first proposed by Penfield and Roberts (1959) and then popularized by Lenneberg (1967) who stated that there are maturational constraints on the time a first language can be acquired. The theory was subsequently extended to L2 acquisition and it has been the subject of the debate ever since. The common belief that the earlier an individual begins L2 learning, the better results they may achieve, seems to be more reasonable in the case of L2 phonology acquisition than other aspects of L2 acquisition in general (Flege, 2007). The age factor seems to be especially influential on segmental aspects of L2 speech production and perception (Flege and MacKay, 2011). It has been proved that during adulthood, most L2 learners perceptually assimilate some non-native speech sounds with similar sounds from their L1 (Best et al., 1988). A lot of research has been done on this topic concerning both L2 learners subjected to formal instruction in L2 and those who acquired their TL in conditions of immersion (e.g. Flege et al., 1999; Flege et al., 2001; Flege and MacKay, 2011). The results of those studies served as a support for the hypothesis that L2 learning is limited by maturational constraints. According to this view, some mechanisms which are needed for successful L1 acquisition work less effectively or may even become inaccessible for use in SLA as humans mature (Flege, 2007).

However, there are some reasons to question the hypothesis. Firstly, other numerous studies show that even adult learners may achieve high competence in the area of L2 speech production. For instance, there is evidence that some adults are highly successful in learning to pronounce L2 words (e.g. Bongaerts, 1999; Golestani and Zatorre 2009). What is more, there are cases of adults who began learning their L2 in childhood, and also children, who produce sentences in their L2 with detectable foreign accents. These differences in L2 production cannot be attributed to the passing of a critical period ending even as early as at the age of 6 (De Keyser, 2000) simply because some foreign-accented early learners started learning their second language prior to that age (Flege, 2007). On the other hand, there is the case of L2 learners acquiring TL via formal instruction in their home country. The common belief that “the earlier an individual begins L2 learning, the better” has been supported by numerous observations of FL teachers in term of L2 pronunciation. However, again numerous studies prove that adult L2 learners are also capable of achieving proper pronunciation or even native-like proficiency in this area. It seems that the difficulties encountered by adult L2 learners may be purely physiological in most cases (Singleton and Ryan, 2004).

4. Current study: rationale and study design

The aim of this study was to examine if elementary, adult Polish learners of German are able to distinguish four German vowels, namely: /ɪ/, /i:/, /ʏ/ and /y:/, organized as two sets of minimal pairs, that is /ɪ/ vs. /ʏ/ and /i:/ vs. /y:/. The vowels were chosen since /y:/ and /ʏ/ (usually represented in spelling as “ü”) are considered to be very difficult sounds for foreigners (Becker, 2012) also for Polish learners (e.g. Bęza, 2001) and are frequently substituted in production by either Polish /i/ or Polish /u/. The choice of sound substitution is usually based on spelling of a given word or on personal preferences (Sobkowiak, 2004).

Twelve elementary, adult (29-52 years old; mean 39; median 38) Polish learners of German agreed to participate in the study. There were three women and nine men. They had been learning German for a year and a half in a language school in the city of Sosnowiec, Poland. None of them had studied German before the course, so they presented a uniform level of proficiency in this language. The subjects had just begun their A2-level language course, however, what was crucial was the fact that they were systematically taught the basics of German pronunciation during their A1-level course.

The study took place in a language laboratory of the language school where the subjects studied German. They were presented a printed list of word pairs differing only in terms of the analysed vowels and listened to the recorded words through headphones. The words were recorded by a native German female speaker who works as a German teacher in Poland. The subjects listened to the playlist twice. There were 6 minimal pairs (plus distracters which were not analysed) printed twice x 12 subjects = 144 analysed instances. The tokens were randomized. The subjects were asked to circle the right option in each pair.

The whole procedure was repeated for the control group which consisted of ten elementary language learners representing a similar level of proficiency in German as an L2. They were 24-47 years old (mean 33; median 33). There were six female and four male subjects. They all started learning German about two years prior to the study. However, those study participants had been taught by another teacher who used a different textbook (which did not contain any elements of pronunciation training) and who did not incorporate any pronunciation module into the course curriculum. In the case of the control group there were 6 minimal pairs (plus distracters which were not analysed) printed twice x 10 subjects = 120 analysed instances.

Table 1
The list of analysed minimal pairs

/ɪ/	/ʏ/
Liste	Lüste
Kiste	Küste
Lifte	Lüfte
/i:/	/y:/
Biene	Bühne
Kiel	kühl
liegen	lügen

The study was inspired by three issues. First of all, the author's earlier pieces of research revealed that even advanced L2 learners frequently were unable to produce new vowels correctly and she wanted to test whether this was not mainly due to the impaired perception of L2 sounds. Another matter was whether the age of learners would affect their L2 segmental

perception. The last question was: to what extent will the systematic training in German pronunciation from the very first contact with the language influence the development of the learners' perception and production in L2?

5. Results

The results revealed that although the subjects from the experimental group were only elementary users of German and started learning the language as adults, they were able to recognize sounds correctly in a considerable number of instances. Naturally, they did not achieve the full correctness in their answers, however they scored much better than it can usually be observed in the case of elementary groups of learners. The table below presents the average results for the analysed sounds.

Table 2
The results obtained by the subjects

Vowel pair	Printed words	Correct words	Correct recognitions
/ɪ/ and /ʏ/	Liste/ Lüste	Lüste	79.2%
/ɪ/ and /ʏ/	Kiste/ Küste	Kiste	75%
/ɪ/ and /ʏ/	Lifte/ Lüfte	Lüfte	83.3%
/i:/ and /y:/	Biene/ Bühne	Bühne	58.3%
/i:/ and /y:/	Kiel/ kühl	Kiel	50%
/i:/ and /y:/	liegen/ lügen	lügen	41.6%

Table 3
The results obtained by the control group

Vowel pair	Printed words	Correct words	Correct recognitions
/ɪ/ and /ʏ/	Liste/ Lüste	Lüste	30%
/ɪ/ and /ʏ/	Kiste/ Küste	Kiste	35%
/ɪ/ and /ʏ/	Lifte/ Lüfte	Lüfte	25%
/i:/ and /y:/	Biene/ Bühne	Bühne	45%
/i:/ and /y:/	Kiel/ kühl	Kiel	35%
/i:/ and /y:/	liegen/ lügen	lügen	40%

As can be seen in the table above, in the case of some of the analysed words, the subjects frequently achieved considerably good results (one must remember that they were elementary users of German who started learning this language as adults).

The results obtained by the experimental group were much better than those obtained by the control group. It can be easily read from the numbers in the tables above, but also the results of the T-test for two independent means. The *p*-value is .002721 which provides us with the statistically significant result.

What is striking is the fact that the number of correct recognitions in the experimental group was much higher for the pair /ɪ/ vs. /ʏ/ than /i:/ vs. /y:/. It is especially surprising in the view of the spacial proximity between the vowels in each pair. The figures below present a model vowel chart for German monophthongs and the actually produced analysed vowels plotted on a vowel plane with the use of Praat 5.3.12 speech-analysis software package (Boersma, 2001). The frequencies of the first two formants are similar to those presented in alike studies (e.g. Pätzold and Simpson, 1997).

Figure 1
German monophthongs (after Mangold, 2005)

Figure 2
German monophthongs actually produced for the purpose of the current study

Table 4
Mean F1 and F2 frequencies of the vowels produced for the purpose of the current study

Vowel	F1 frequencies (Hz)	F2 frequencies (Hz)
/ɪ/	390	2135
/ʏ/	430	2100
/i:/	330	2350
/y:/	335	2200

As can be seen in Figure 1, Figure 2 and Table 3, the proximity between /ɪ/ and /ʏ/ is smaller than between /i:/ and /y:/. The computed Euclidean distance between /ɪ/ and /ʏ/ equals =53 Hz while for /i:/ and /y:/ it is =150 Hz. Because of that, one could expect subjects to encounter greater difficulties in distinguishing between /ɪ/ and /ʏ/ which are characterized by similar acoustic properties, however the results were nothing but the opposite. The best recognized pair was “Lifte/ Lüfte” while the worst results were obtained for the pair “liegen/ lügen”. The total results for each pair of vowels are as follows:

Figure 3
Total correct and incorrect recognitions of words containing the /ɪ/ and /ʏ/ vowels (the experimental group)

Figure 4
Total correct and incorrect recognitions of words containing the /i:/ and /y:/ vowels (the experimental group)

It is interesting that the results obtained by the control group were quite the opposite. In this case, although the general number of correct sound recognitions was significantly lower, the subjects could better distinguish between /i:/ and /y:/ (which is understandable since the two sounds are less “similar” to each other in terms of their acoustic properties than /ɪ/ and /ʏ/). This may also suggest that the phonetic training sensitizes language learners to even slight changes in the acoustic signal they hear.

Figure 5

Total correct and incorrect recognitions of words containing the /ɪ/ and /ç/ vowels (the control group)

Figure 6

Total correct and incorrect recognitions of words containing the /i:/ and /y:/ vowels (the control group)

6. Conclusions

Although the number of subjects was relatively small and the study should be treated as preliminary, the results described above contribute to the statement that it is possible to train second (or: foreign) language learners in the field of L2 phonetics and phonology so that they are able to recognize most of the foreign sounds (and thus words) correctly. Even those language users who started learning a new language as adults, may achieve relatively good results and their perception can be trained to a significant extent. What is more, starting phonetic training as early as at the beginner (A1) level can provide learners with noticeable benefits. Correct perception and pronunciation facilitate mutual understanding between

speakers and listeners and thus make the whole process of communication much easier. It is especially important for beginner and elementary language learners who start their communication with other L2 users and their initial successes and failures may significantly affect their attitudes towards further language learning. Naturally, at the beginning not all sounds will be differentiated from the others, however this skill can be trained with the further practice. But, as one of the subjects noticed, “it is never too late to start learning and never too late to spot a difference between the things you did not know anything about”.

7. References

- Becker, T. (2012). *Einführung in die Phonetik und Phonologie des Deutschen*. Darmstadt: WBG.
- Best, C. T. (1994). The Emergence of Native-Language Phonological Influences in Infants: A Perceptual Assimilation Model. In Goodman, C. and Nusbaum, H. (eds.), *The Development of Speech Perception*. Cambridge: The MIT Press: 167-224.
- 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*. Timonium: York Press: 171-204.
- Best, C.T., McRoberts, G.W. and Sithole, N.M. (1988). Examination of Perceptual Reorganization for Nonnative Speech Contrasts: Zulu Click Discrimination by English-speaking Adults and Infants. In: *Journal of Experimental Psychology: Human Perception and Performance* 14 (3).
- Bęza, S. (2001). *ABC der deutschen Phonetik*. Warsaw: Poltext
- Boersma, P. (2001). Praat, a System for Doing Phonetics by Computer. *Glott International* 10: 341-345.
- Bongaerts, T. (1999). Ultimate Attainment in L2 Pronunciation: The Case of Very Advanced Late Learners. In: Birdsong, D. (ed.) *Second Language Acquisition and the Critical Period Hypothesis*. Mahwah NJ: Lawrence Erlbaum Associates.
- Brown C. (2000). The Interrelation Between Speech Perception and Phonological Acquisition from Infant to Adult. In: Archibald, J. (ed.) *Second Language Acquisition and Linguistic Theory*. Malden: Blackwell Publishers Inc.
- De Keyser, R. (2000). The Robustness of Critical Period Effects in Second Language Acquisition. In: *Studies in Second Language Acquisition* 22: 499-534.
- Eckman, F. (1977). Markedness and the Contrastive Analysis Hypothesis. In: *Language Learning*, 27, 315-330.
- Eddine, A.N. (2011). Second Language Acquisition: The Articulation of Vowels and Importance of Tools in the Learning Process. In: Arabski, J. and Wojtaszek, A. (eds.) *The Acquisition of L2 Phonology*. Bristol: Multilingual Matters.
- Flege, J. (1991). Orthographic Evidence for the Perceptual Identification of Vowels in Spanish and English. In: *Quarterly Journal of Experimental Psychology* 43: 701-731.
- Flege, J.E. (1992). The Intelligibility of English Vowels Spoken by British and Dutch Talkers. In: Kent, R. (ed.) *Intelligibility in Speech Disorders: Theory, Measurement and Management*. Amsterdam: John Benjamins.
- Flege, J.E. (1995). Second Language Speech Learning: Theory, Findings and Problems. In: Strange, W. (ed.) *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Timonium: York Press.
- Flege, J.E. (2007). Language Contact in Bilingualism: Phonetic System Interactions. In: Cole, J. and Hualde, J.I. (eds.) *Laboratory Phonology* 9. New York: Mouton de Gruyter.
- Flege, J.E., Bohn O-S. and Jang, S. (1997). Effects of Experience on Non-Native Speakers' Production and Perception of English Vowels. In: *Journal of Phonetics* 25, 437-470.

- Flege, J.E., Yeni-Komshian, G. and Liu, S. (1999). Age Constraints on Second Language Acquisition. In: *Journal of Memory and Language* 41:78 – 104.
- Flege, J.E. and MacKay, I. (2011). What Accounts for “Age” Effects on Overall Degree of Foreign Accent? In Wrembel, M. Kul, M. and Dziubalska-Kołaczyk, K. (eds.) *Achievements and Perspectives in the Acquisition of Second Language Speech: New Sounds 2010*, Vol. 2, Bern, Switzerland: Peter Lang.
- Golestani, N. and Zatorre, R.J. (2009). Individual Differences in the Acquisition of Second Language Phonology. In: *Brain and Language* 109 (2009) 55-67.
- Lenneberg, E.H. (1967). *Biological Foundations of Language*. New York: Wiley and Sons.
- Lipińska, D. (2013a). ‘Luke, These Sheep’s Seen Keen’ (and Other Misperceptions of a Learner’s Ear). In: *Proceedings of 3rd International Conference on English Pronunciation: Issues & Practices* (pp. 69-72). CD-ROM version
- Lipińska, D. (2013b). Influence of Formal Instruction in English Phonetics and Phonology on Polish learners’ Production of English /æ/. In: Przedlacka, J., Maidment, J. and Ashby, M. (eds.) *Proceedings of PTLC 2013. Papers from the Phonetics Teaching and Learning Conference, London, 8–10 August 2013* (pp. 59-62). London, Phonetics Teaching and Learning Conference Chandler House (e-book PDF)
- Littlewood, W.T. (1994). *Foreign and Second Language Learning : Language-Acquisition Research and its Implications for the Classroom*. Cambridge: Cambridge University Press.
- Mangold, M. (2005). *Duden Band 6: Das Aussprachewörterbuch*. Mannheim: Dudenverlag.
- Nowacka, M. (2010). The Ultimate Attainment of English Pronunciation by Polish College Students: a Longitudinal Study. In: E. Waniek-Klimczak (ed.) *Issues in accents of English 2. Variability and Norm*. Newcastle: Cambridge Scholars Publishing.
- Pätzold, M., and Simpson, A. P. (1997). Acoustic Analysis of German Vowels in the Kiel Corpus of Read Speech. *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung Universität Kiel*, 32, 215-247.
- Penfield, W. and Roberts, L. (1959). *Speech and Brain Mechanisms*. Princeton: Princeton University Press.
- Pilus, Z. (2005). Perception of Voicing in English Word-Final Obstruents by Malay Speakers of English: Examining the Perceptual Assimilation Model. In: Wijasuriya, B.S. (ed.) *Malaysian Journal of ELT Research* 1.
- Porzuczek, A. (2007). English Vowel Quantity in Polish Learner’s Speech Perception and Production. In: J. Arabski, D. Gabryś-Barker and A. Łyda (eds.) *PASE Papers 2007: Studies in Language and Methodology of Teaching Foreign Languages*. Katowice: PARA: 96-105.
- Rojczyk, A. (2009). Modele Percepcji Systemu Dźwiękowego Języka Obcego. In: J. Nijakowska (ed.) *Język Poznanie Zachowanie: Perspektywy i Wyzwania w Studiach nad Przyswajaniem Języka Obcego*. Łódź: Wydawnictwo Uniwersytetu Łódzkiego: 120-135.
- Rojczyk, A. (2010a). Forming New Vowel Categories in Second Language Speech: The Case of Polish Learners’ Production Of English /ɪ/ and /e/. In: *Research in Language* 8: 85-97.
- Rojczyk, A. (2010b). Production and Perception of Vowel /æ/ by Polish Learners of English. In: K. Dziubalska-Kołaczyk, M. Wrembel, M. Kul (eds.) *Proceedings from the Sixth International Symposium on the Acquisition of Second Language Speech* [CD ROM version].
- Singleton, D. And Ryan, L. (2004). *Language Acquisition: The Age Factor. 2nd Edition*. Clevedon: Multilingual Matters Ltd.
- Sobkowiak, W. (2004). *English Phonetics for Poles*. Poznań: Wydawnictwo Poznańskie.