PHONETIC TRANSCRIPTION OF ATYPICAL SPEECH: SECOND OR A FOREIGN LANGUAGE

Ines Carović

Department of Phonetics
Faculty of Humanities and Social Sciences
University of Zagreb
icarovic@ffzg.hr

ABSTRACT

The paper shows the results of significant research to "create a large set of new speech sound categories and to expand the listener's perceptual system to detect differences between that expanded set of sounds" through training [1]. The research used phonetic transcriptions of 360 first year BA students in Phonetic Studies after 16 and after 60 hours of training. The students had to transcribe a foreign language speaker and results of the study show that there is inadequate time for correct transcription. In addition, a questionnaire given to students show that it was easier to write broad transcription rather than narrow; but when it came to the accuracy of pronunciation or being able to hear and understand, it was more helpful to write narrow transcription. Also, vowels were more difficult to transcribe than consonants. The research showed that the multilingual students were better in phonetic transcription than their monolingual peers.

Keywords: phonetic transcription, transcription training, perception of foreign language

1. INTRODUCTION

Lee [2] analysed articles and research on the importance of listener training for perceptual evaluation of speech. [1, 3, 4, 5, 6, 7, 8] which dealt with the training of speech perception and transcription. They all agree with Stemberger & Bernhardt [1] that it is most important to "create a large set of new speech sound categories and to expand the listener's perceptual system to detect differences between that expanded set of sounds" through training and that both monolingual and multilingual contexts are important for perception training, as they expand perceptual sensitivity to fine distinguishing of phonetic nuances. However, Müller & Papakyritsis [5] note that we should "not let a 'segmental mindset' become a tyrant dominating our thinking about speech, whether normal or disordered". The paper researched how students listen to, transcribe, and pronounce specific speech and the differences after training in phonetic transcription

(PT).Some authors believe that the best transcription comes with video and audio recordings [1, 3] and supplemented with instrumental techniques [1, 9], such as the ultrasound [10, 11] and the spectrogram [12]. Although after the completion of lectures in articulatory and acoustic phonetics, the students might use spectrograms based on acoustic recordings, this was not applied in the present research. Only two students stated in the questionnaire that in the recording of the Chinese speaker they wanted to see whether the spectrograms showed that all voiceless phonemes were indeed voiceless, and they should be voiced in Croatian language. Regarding the time spent on PT training, different authors state different times required for high-quality PT: "100 hours … received by some European speech and language therapist (SLT) students is probably insufficient" [3], there should be more practice [6], "consistent practice" [8], training spread over the years [1, 8], and some mentioned the need of lifelong learning – Refresher Course for Researchers & Clinicians [1, 7]. The most important goal of the course, but also of the study of phonetics is "accuracy, being able to hear and understand, knowledge of symbols for speech sounds" [12]. The authors dealt with the effect of concurrent speech disorders on auditory-perceptual evaluation: effect of pitch, vocal intensity, breathiness, hoarseness (or roughness), articulation on judgements of hypernasality; effect of hypernasality on judgements of articulation, intonation, effect of pitch, vocal intensity on judgements of breathiness, harshness, hoarseness [14]; however, this research omits all labels for pitch, intensity, supralaryngeal placements, suprasegmental features – unless this was extensively noted. Nelson et al [7] mention a broad and narrow PT ("Broad – recording of phonemes which is language dependent, does not contain detail"; "Narrow – detail of pronunciation is recorded"). A question arises: in PT training, how difficult is it for students to perceptually notice the differences between a language which is their mother tongue and the pronunciation of that language as a second or foreign one? The assumption is that after the training it will be easier for them to choose the symbol for perceptual difference although, phonemically, that phoneme is a part of realisation, and this is the precise reason why they will more easily opt for narrow PT

when they are asked to repeat identical pronunciation as well. With phonemes and differences not in their language system or the system of the language they learned, they will find it difficult to agree on the PT symbol with their partner. Main research question for this study was how good are students results of PT after 60 hours of training. The study goal is to compare this training results with results from different areas of research [1, 3, 6, 7, 8].

2. RESEARCH

The research analysed 360 transcribed texts based on 22 audio recordings of foreign student speeches who were learning Croatian and who had a different levels of Croatian language knowledge. Transcriptions were annotated in pairs by 1st year undergraduate students (average age of 19,7) of enrolled in a Phonetic Transcription Coursein the Department of Phonetics, Faculty of Humanities and Social Sciences, in the 4th and 15th week of the semester, , after 16 and 60 hours of training, respectively. The same students received different audio recordings of speakers twice. Mother tongues of speakers were : English, Chinese, Japanese, Vietnamese, Spanish, Arab, Danish, Hungarian, German, Malay, Bulgarian, Russian, Turkish, Catalan, Portuguese, Slovenian and Italian. Some languages had speakers from different areas. For example, some English speakers were from USA, South Africa, and Australia, while Spanish speakers were from Chile, Spain, and Argentina. All speakers read the fable by Aesop, The North Wind and the Sun [15], which was very familiar in Croatian to the students. The students were advised to listen to the audio recording with their headphones on and not to focus on the voice quality, but only on consonants and vowels. The instruction was to agree in pairs on transcription symbols and use the symbols to read the text as the speaker they analysed. At the beginning of the semester, the students received 16 hours of introduction to PT and transcription practice on the example of Croatian standard language (typical speech), as well as the perceptual speech analysis, annotation of transcription symbols of broad PT for vowels and consonants in Czech and Catalan.

At the end of the semester, after 60 training hours in PT, the students transcribed in pairs the new acoustic example. Following both transcriptions, the subjects presented their results in a group setting, received a model PT to check the transcription of the recording they analysed and received feedback of the whole group on whether they agree with the transcription. After the phonetic transcription, the students/subjects received a questionnaire on their impressions and the transcription process.

The analysis of the transcribed texts was made in relation to the transcription model made by expert phoneticians.

3. RESULTS

3.1. Questionnaire

In 97% of cases, respondents felt that PT was easier at the end of semester, but their comments from the questionnaire indicate that at the end of semester they are aware they are not accurate. Respondents are also not confident they are accurate in PT. Students responded that it was easier to write broad transcription rather than narrow, but when it came to the accuracy of pronunciation or being able to hear and understand, it was more helpful to write narrow transcription. Also, vowels were more difficult to transcribe than consonants. Table 1 shows answers to questionnaire questions after each PT. The scale for answers was from 1 to 5. 1 – I do not agree at all; 5 – I completely agree.

Question	Results after the 1st PT	Results after the 2nd PT
1. We agreed easily in pair on the PT symbol.	3.5	3.7
2. I am completely certain regarding the PT.	2.1	3.0
3. It is easy to transcribe the broad PT.	3.6	3.2
4. It is easy to transcribe the narrow PT.	3.3	3.6
5. It is useful to transcribe the broad PT.	3.0	3.5
6. It is useful to transcribe the narrow PT.	4.1	4.5
7. I am certain regarding the consonant PT.	3	3.9
8. I am certain regarding the vowel PT.	2.2	2.4
9. I have studied the mother tongue of the speaker for PT.	2.1	1.9
10. I have listened to the recording more than 100 times.	4.4	4.6

Table 1: Answers to the questionnaire. The scale for answers was from 1 to 5. 1 – I do not agree at all; 5 – I completely agree.

3.2. Transcription accuracy at the beginning and at the end of the semester

Although the subjects did not learn most of speakers' mother tongues, the number of deviations and potential changes in the PT in relation to Croatian standard language was very different with different speakers, so a percentage of accurate PTs in relation

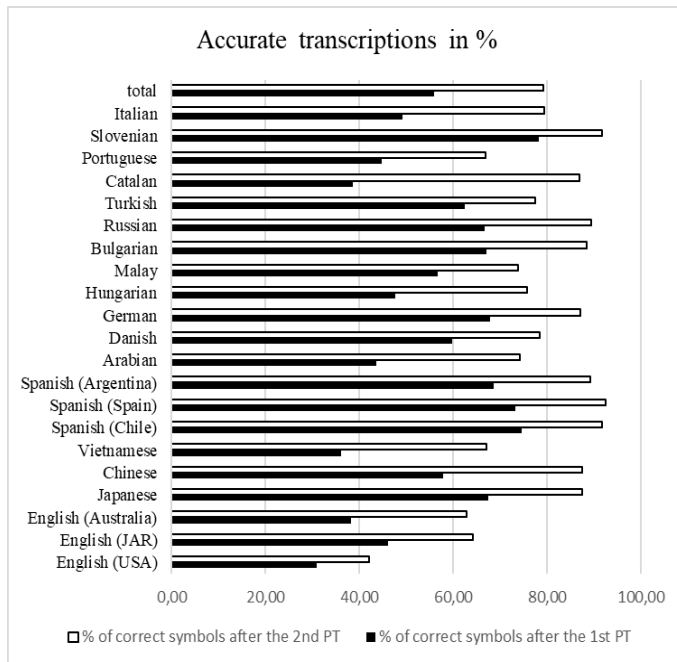to the overall number of differences in comparison with standard Croatian was calculated, as shown in Chart 1.

Chart 1: The percentage (%) of accurate transcriptions (that is, changes in relation to the Croatian text).

The subjects made fewer errors in the second PT, and Russian and Spanish were the languages that had fewest differences between the model and the subjects.

It is notable that the percentage of correct PT symbols increased in the second PT after 60 hours of training, to 79,2%, however, it is still not enough for a certain PT. We can conclude that the students would definitely need more practice and training, especially for atypical speeches or speeches they have not been in contact with.

In the second PT, the students who evaluated speakers of Spanish, Slovenian and Russian as mother tongue, performed the best transcription; the reason for this could be that Slovenian and Russian are Slavic languages, so they are closer to the mother tongue of subjects, while Spanish is a language that does not have a large number of deviations in the phonemic system from the Croatian language. Also, although the subjects did not learn it within an institution, Spanish was a language they were most exposed to at an earlier age, through the television. The success of English speakers' PT is surprisingly low. After the analysis, it was established that the level of accuracy was most influenced by the level of learning Croatian as a foreign language, regardless of the fact that these languages were closer to the students as phonemic systems.

The subjects estimated they had more confidence in PT at the end, rather than at the beginning of the semester, but that they cannot be 100% sure for some transcribed elements of the speech. The results of the analysis show that there are no 100% correct answers in PT, but that incorrect answers pose a problem to most listeners as well after the presentation of the results. Figures 1 and 2 show PT examples for speakers of English and Russian as mother tongue in the end of the semester.

Figure 1: Transcription example of a speaker of English as mother tongue.

Figure 2: Transcription example of a speaker of Russian as mother tongue.

Some "incorrect" transcriptions were annotated as such as they were annotated in too much detail in the narrow PT. Therefore, it is important to be aware of the agreement regarding the final decision on the PT and the explanation for future collaborators as well. Howard & Heselwood [3] state that "there are no right answers, only more or less likely answers". Agreement of transcriptions: "fail to produce an accurate reflection of the similarities between listeners' perceptions precisely because they rely on scores of symbol agreement, rather than symbol equivalence" [3], and "do not reflect the range of discrepancies that might occur between two transcribers" [4].

3.3. Students who knew the mother tongue of speakers who learned Croatian as a second or foreign language

The fewest deviations from the model PT (phonetic expert) were expected from students who learned the mother tongue of speakers from the recording, and especially the students who studied that language. Given that they were familiar with the phonetic features of the language, and they wrote in the questionnaires that it was easier to "listen" to them as they could transcribe with more certainty what they heard, this was not completely confirmed. However, it was confirmed that the students who knew more than one language had more success in the PT. The most likely reason for this is a more developed phonological awareness and perception for fine differences. Students who did not know the mother tongue wrote that they had to listen many times to the recording to be certain of they heard, but that after many instances of listening they would more easily create a repetitive system, that is, a system of errors in foreign language. It is difficult to state the causes for some errors in PT made by students who knew the mother tongue of the speaker they analysed. Some authors [3, 4] state that in phonetic transcription "predictions of expected speech sounds through knowing the intended target form also have a part to play in producing unreliable transcriptions".

4. CONCLUSION

We can conclude from this research that after the PT training, students become better with perception, transcription symbol annotation and the pronunciation of these symbols. However, 60 hours of training are not enough for obtaining a unified group of reliable listeners. Howard and Heselwood have mentioned 100 hours of training which "is probably insufficient" [3], Knight [6] have noticed there should be more practice and Shaw & Yanushevskaya [8] talk about consistent practice and training spread over the years as well as Stemberger & Bernhardt [8]. Phonetic transcription, articulatory and acoustic phonetics and more importantly - practice and exercises are good way to improve results in PT. Howard & Heselwood [3] talk about the linguistic background: "a native speaker- hearer's perceptual system will constrain perception of speech data", and [3, 16] agree that "listeners tend to force the new material through the perceptual grid of the phonological categories of their own language". Therefore, it is important to expose the subject to a larger number of different stimuli to develop a bigger sensitivity to small differences, and to categories they are familiar with in their own linguistic system. The present research has confirmed this. As Lee [2] concludes in her work, "perceptual evaluation is the gold standard for assessing speech. We need effective listener training programmes for professional training and continuous professional development". The program at the Department of Phonetics at the University of Zagreb does not yet have research on the PT of atypical speech, although graduate students do practice with annotating and analysing atypical speech. PT is a start for perceptual analysis, error systems in a foreign language and phonetic correction of speech, analysis of developmental children's speech, annotation of typical and atypical speech, and it is necessary if someone wants to do phonetics. Lee [2] concludes that "it is necessary to develop a training programme that suits the curriculum and further investigate the impact of different factors on perceptual evaluation for improving training programmes". Results in this research indicate that 60 hours is too little time for making transcription perfect. Research like this is only an indicator of the need for a systematic training in PT, perceptual evaluation of speech, and perceptual judgements.

5. REFERENCES

[1] Stemberger, J. P., Bernhardt, B. M. (2020). Phonetic transcription for speech-language pathology in the 21st century. *Folia Phoniatrica et Logopaedica, 72*(2), 75-83.
[2] Lee, A. 2022. Listener Training for Perceptual Evaluation of Speech. Plenar talk presented at the Speech Research 2022, 8-10 December 2022, Zagreb, Croatia.
[3] Howard, S. J., Heselwood, B. C. (2002). Learning and teaching phonetic transcription for clinical purposes. *Clinical Linguistics & Phonetics, 16*(5), 371-401.
[4] Ball, M., Müller, N., Klopfenstein, M., Rutter, B. (2009). The importance of narrow phonetic transcription for highly unintelligible speech: Some examples. *Logopedics Phoniatrics Vocology, 34*(2), 84-90.
[5] Müller, N., Papakyritsis, I. (2011). Segments, letters and gestures: Thoughts on doing and teaching phonetics

and transcription. *Clinical Linguistics & Phonetics, 25*(11- 12), 949-955.

[6] Knight, R. A. (2010). Sounds for study: Speech and language therapy students' use and perception of exercise podcasts for phonetics. *International Journal of Teaching and Learning in Higher Education, 22*(3), 269-276. https://files.eric.ed.gov/fulltext/EJ938562.pdf

[7] Nelson, T. L., Mok, Z., Ttofari Eecen, K. (2020). Use of transcription when assessing children's speech: Australian speech-language pathologists' practices, challenges, and facilitators. *Folia Phoniatrica et Logopaedica, 72*(2), 131-142.

[8] Shaw, A., Yanushevskaya, I. (2022). Students' views and experiences of the training and use of phonetic transcription in speech and language therapy – The Irish perspective. *Clinical Linguistics & Phonetics, 36*(2-3), 276-291.

[9] Shriberg, L. D., Lof, G. L. (1991). Reliability studies in broad and narrow phonetic transcription. *Clinical*

[10] *Linguistics & Phonetics, 5*(3), 225-279.

[11] Cleland, J., Lloyd, S., Campbell, L., Crampin, L., Palo, J. P., Sugden, E., Zharkova, N. (2020). The impact of real-time articulatory information on phonetic transcription: Ultrasound-aided transcription in cleft lip and palate speech. *Folia Phoniatrica et Logopaedica, 72*(2), 120-130.

[12] Sugden, E., Cleland, J. (2022). Using ultrasound tongue imaging to support the phonetic transcription of childhood speech sound disorders. *Clinical Linguistics & Phonetics, 36*(12), 1047-1066.

[13] Sullivan, K., Czigler, P. (2002). Maximising the educational affordances of a technology supported learning environment for introductory undergraduate phonetics. *British Journal of Educational Technology, 33*(3), 333-343. https://doi.org/10.1111/1467-8535.00268

[14] Speights Atkins, M., Bailey, D. J., Seals, C. D. (2022). Implementation of an automated grading tool for phonetic transcription training. *Clinical Linguistics & Phonetics*. Advance online publication. https://doi.org/10.1080/02699206.2022.2048314

[15] Lee, A., Potts, S., Bressmann, T. (2020). Speech- language therapy students' auditory-perceptual judgements of simulated concurrent hypernasality and articulation disorders. *Clinical Linguistics & Phonetics, 34*, 479- 492.

[16] International Phonetic Association. 2009. *Handbook of the International Phonetic Association*. London: Cambridge University Press.

[17] Laver, J. (1991). *The gift of speech: Papers in the analysis of speech and voice*. Edinburgh University Press.