Elisa REUNANEN, Antti SALORANTA,
Maija S. PELTOLA, Alan WRENCH

THE ARTICULATORY DESCRIPTION OF FINNISH VOWELS USING ULTRASOUND

Abstract

Speech has mostly been studied using measurements of speech sound acoustics. Frequencies of speech are an indirect measure of articulatory movements, and the description of articulatory has mostly been based on tactile and visual feedback of tongue and lips. The aim of the present study was to describe the articulatory system of Finnish vowels using ultrasound. Subjects’ speech data was collected during a production task in which subject read aloud pseudowords created according to Finnish phonotactic rules. Articulatory movements of the target vowels in words were analysed using Articulatory Assistant Advanced software. Vowel articulatory system was consistent and there was no great within-subject variation. Anatomical differences cause between-subject variation, but despite that, the vowel articulation form a system in which the articulatory movements of different vowels are in same relation with each other. In addition, we noticed that rounding the lips is not the only articulatory movement differentiating rounded and unrounded front vowels in Finnish.

Keywords: *vowel articulation, Finnish, articulatory movements, ultrasound*

Somu valodas patskaņu artikulārais raksturojums, izmantojot ultraskaņu

Kopsavilkums

Runa galvenokārt tiek pētīta, izmantojot skaņu akustiskos mērījumus. Runas skaņu frekvences netieši norāda uz artikulatoru kustībām, un artikulācijas apraksti galvenokārt tiek balstīti uz taktilajām un vizuālajām pazīmēm, kas liecina par mēles un lūpu kustībām. Šī pētījuma mērķis ir aprakstīt somu valodas patskaņu artikulāro sistēmu, izmantojot ultraskaņu.

Ieraksta laikā informanti nolasīja preidovārdus, kas tika konstruēti saskaņā ar somu valodas fonotaktikas likumiem. Mērķa patskaņiem atbilstošās artikulārās kustības vārdos tika analizētas, izmantojot datorprogrammu *Articulatory Assistant*. Patskaņu artikulārā sistēma bija konsekventa, un nozīmīgs starpsubjektu mainīgums netika novērots. Anatomiskas atšķirības veicina starpsubjektu mainīgumu, taču, tam par spīti, patskaņu artikulācija veido sistēmu, kurā dažādiem patskaņiem atbilstošās artikulārās kustības atrodas līdzīgās savstarpējās attieksmēs. Pētījumā arī novērots, ka lūpu noapaļojums nav vienīgā artikulārā kustība, kas atšķir noapaļotos un nenoapaļotos patskaņus somu valodā.

Atslēgvārdi: *patskaņu artikulācija, somu valoda, artikulārās kustības, ultraskaņa*

1. Introduction

Vowel articulation has mostly been studied using measurements of speech sound acoustics, which is in fact an indirect measure of articulation. By studying acoustics, it is only possible to study the outcome of articulation and infer the gestures behind it. In addition to acoustic measurements, articulatory motor coordination has been studied using X-ray photographs in earlier studies (Sovijärvi 1967; Johansson et al. 1982, Wood 1982). These photographs, however, give only a static image of the position of the vocal organs. Magnetic resonance imaging (MRI), palatography and electromyography have also been used in articulation studies. MRI allows the studying of movements, but the quality of images and noise during registration cause problems in analysis. Palatography gives indirect information about muscles, e.g. which parts of the tongue have touched the palate. Electromyography evaluates and records the electrical activity produced by skeletal muscles. In palatography and electromyography, research instruments must be placed in or on the subject's mouth, which may affect production. (Proctor et al. 2015, Hoole 2006, Anderson et al. 2004.) Ultrasound is the newest research method in speech production studies. It has mostly been used for studying the physical mechanism of speech production and speech therapy, rather than for language-specific studies. (Cleland et al. 2016, Wrench et al. 2011.)

The acoustics of a vowel are the outcome of two factors: the source of sound and the resonant system, which are at least partly independent from each other. The breathing mechanism – the lungs and the muscles in the chest and abdomen – works as an energy supply for speech. Air forced out of the lungs causes the vocal cords to vibrate, which works as a source of sound for speech. The resonant system, through which the air flows, consists of everything above the larynx. Articulatory movements cause changes in the vocal tract, which as a resonant system affects the frequencies that differentiate sounds. (Suomi 1990: 79–80; Fry 1980: 61–62, 71; Ladefoged 1962: 89–90; Jones 1962: 15–16.) As is well known, the most important factors in vowel differentiation are the frequencies of the first and second formant (F1 and F2), which can be used to differentiate all vowels from each other. Changes in formants are the results of different articulatory movements, which affect the shape and the dimensions of the vocal tract. Conventionally, vowels are classified into front, central and back vowels according to the position of the main part of the tongue, into close, close-mid, open-mid and open vowels according to the height to which the tongue is raised and into rounded and unrounded vowels according to the rounding of the lips. The relation between articulation and the formants has certain regularities: Rounding of the lips decreases the frequencies of F2 and F3. When the tongue moves to a lower position, the frequency of F1 rises and when the tongue moves forward, the frequency of F2 rises. Tongue is the most important organ for shaping the vocal tract, because the tract itself is quite inflexible. Bones set their own boundaries to movements and the angle towards larynx is always the same. (Fry 1980: 76; Joos 1948: 50, 57–59; Ladefoged 2001: 39, 1975: 173; Jones 1962: 16–17; 19.) The relation between articulation and acoustics is not always rectilinear, and it is possible that a minor shift in articulation can cause considerable difference in acoustics – and sometimes a major shift in articulation does not cause a significant difference in acoustics. (Stevens 1972).

The Finnish vowel system includes eight vowels: /ɑ, e, i, o, u, y, æ, ø/. Mid vowels in Finnish are actually approximately half-way between the IPA close-mid and open-mid vowels, however, in this paper, the IPA cardinal vowel symbols without diacritics will be used for simplicity. All Finnish vowels have short and long variants, and the difference between sounds is significant for the sound system. Earlier acoustic studies on Finnish vowels (Suomi 1990; Suomi et al. 2008) have presented the average formant values for all vowels and it has been shown that longer vowel more extreme values and when presented in a vowel diagram, they show a hyperspace quality (O'Dell 2013; Nakai et al. 2012; Nakai et al. 2015.)

The purpose of this study is to describe the articulatory motor coordination movements of the tongue and the lips in Finnish vowels and to study whether the articulation in accordance with descriptions in theory: what are the properties of the vocal tract, which cause the changes in formants. In addition, the aim was to test the ultrasound system as a research method of speech articulation.

2. Methods

Ultrasound is sound waves, which have a wave length above the human perceptual threshold, from 20 kHz up to several gigahertz. It can be used for detecting objects and for measuring distances. Ultrasound waves with different frequencies reflect from different materials in different ways. Frequencies of 2 MHz and higher are used in medical ultrasonic imaging. Ultrasound pulses are sent into tissue using a probe, which converts electrical signal to ultrasound waves and vice versa. The short wavelength of the high frequencies allows resolution of internal details in tissues. Ultrasound is a non-invasive method and causes no harm for subjects. (Grönroos 2010: 22–33.)

2.1. Subjects

Eleven monolingual Finnish women participated in a production task. Participants were 24–40 years old and their mean age was 30 years. They all lived in the Turku area, in South-Western Finland. None of them had any diagnosed speech disorders.

2.2. Stimuli

In the production task, the participants read aloud pseudowords, which were presented on the screen. Pseudowords were used to con-

firm the words were not idiosyncratic and to enable the creation of a series of word forms in which the words differ in the quality of the long vowel only. The dental plosive /t/ was selected as the consonant context because changes in formants within the vowel and the consonant are as small and consistent as possible in a dental context. (Fant 1968: 257–258.) The words were in accordance with Finnish phonotactic rules and each word contained one target vowel, thus resulting in eight targets according to the Finnish vowel system. The words were /tɑ:te, te:te, ti:te, to:te, tu:te, ty:te, tæ:te, tø:te/; the target vowel was the long vowel at the first syllable. (/te:te/ is a lexical word in Finnish. It is a derivate form the verb *teettää* and has a low frequency.) Each word was presented three times, resulting in a total of 24 productions. Long vowels were used to make sure the target vowel had a clear steady-state position. It is worth noticing that short and long Finnish vowels differ in acoustics (O'Dell 2013).

2.3. Procedure

During the production task, subjects' speech data was collected using the Echo Blaster 128 CEXT-1Z ultrasound system to measure articulatory movements during speech production. Ultrasound data was recorded at a frame rate 68fps at a 123,8 degree field of view (FoV). The depth setting was 90mm. The ultrasound recordings were of the mid-sagittal view.

When studying speech, the ultrasound probe is stabilized under the chin with a headset. It requires no mouth internal parts, so speech production is more natural than, for example, with palatography. If the headset and the probe are placed correctly, they do not restrict movements of the chin. (Scobbie et al. 2008.) If the subject wears the headset for a long time, it may feel uncomfortable and heavy. Therefore it is desirable to plan the experiment so that it does not take more than an hour. (Cleland et al. 2015)

The headset was fitted to subjects' heads to stabilize the ultrasound probe. A headset-mounted micro-camera was used to monitor movements of the lips from the front of the subject. Recordings were performed in a quiet laboratory and the full experiment took about 30 minutes. Subjects sat on a chair during experiment.

2.4. Analysis

The ultrasound data was analysed by using Articulate Assistant Advanced (AAA, version 2.16.12) software. In every word, the steady-state point was selected from the ultrasound video. Then a spline indicating the tongue surface was autotracked to the ultrasound video and corrected manually if needed. Every tongue spline from the vowel's steady-state point was taken to AAA's spline workspace. At the spline workspace, the average tongue contours were calculated for each vowel and the eight different vowels proportioned to each other. Due to the different sizes of the subjects' heads and the different shapes of the subjects' mouths, it is not possible to directly compare data across subjects. Therefore every subject was analysed as an individual to see the shape of the vowel articulatory system. The systems of eight vowels were then compared between subjects. Videos from the lip camera were analysed separately with AAA and then compared to tongue splines to study the interaction between tongue and lip articulations. Due to technical difficulties, the lip camera did not record every time, and so it is not possible to analyse lip articulation comprehensively with every vowel.

Two subjects had to be excluded from the data due to poor quality of the ultrasound data, which was the consequence of weak settings at the software during registration. The depth of the ultrasound imaging was not suitable and the field of view was too narrow, consequently, the tongue of the subject could not be seen all the time.

3. Results

3.1. Individual vowels

As can be seen from the Figures 1–8, articulatory motor coordinations are very consistent: all repetitions from one subjects (KH108), and every single vowel and the average tongue position of three same vowel repetitions are presented. At the steady-state position, the shape of tongue is similar every three times the same vowel is produced. This makes it possible to use average tongue surface splines when comparing different vowels of one individual. Within-subject variation is minimal as can be observed from Figures 1–8.

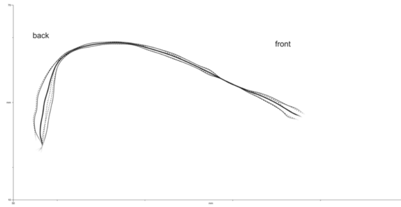

Figure 1. KH108 /u/

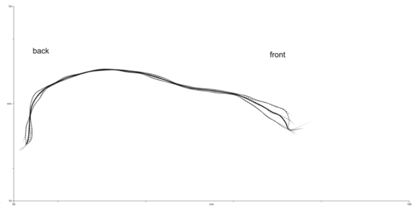

Figure 2. KH108 /o/

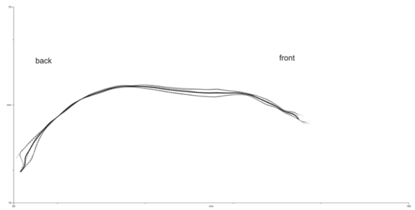

Figure 3. KH108 /ɑ/

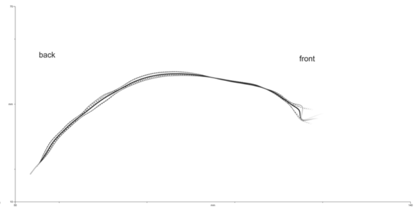

Figure 4. KH108 /æ/

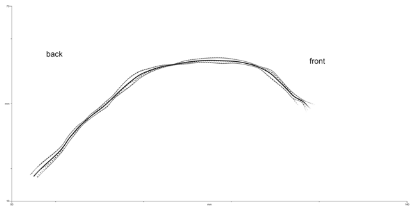

Figure 5. KH108 /ø/

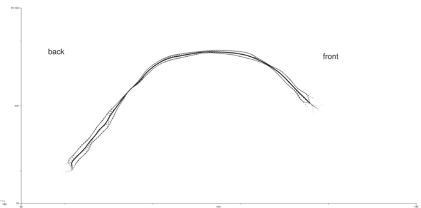

Figure 6. KH108 /e/

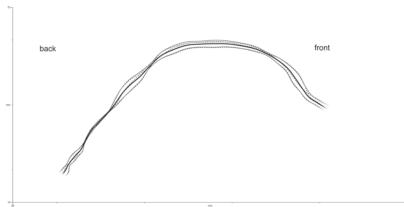

Figure 7. KH108 /y/

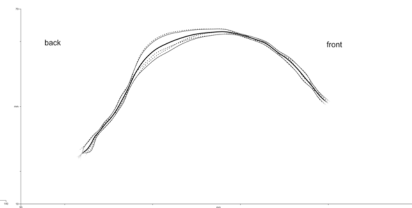

Figure 8. KH108 /i/

Figures 1–8. Solid line represents the average articulatory movement of the vowel. Dotted lines represents three different repetitions of the vowel.

3.2. Vowel articulation as a system

Figures 9–11 contain average tongue surface splines of all eight vowels from three subjects (KH102, KH108, KH112). As becomes evident, the vowel articulatory movements work as a system: The relation between different vowel articulations shows no great between subject variation, hence each subject is able to maintain a solid system. It is obvious that the dimensions of the movements differ between subjects, due to anatomical differences in size. The difference can be seen between front and back vowels and also between close, mid and open vowels.

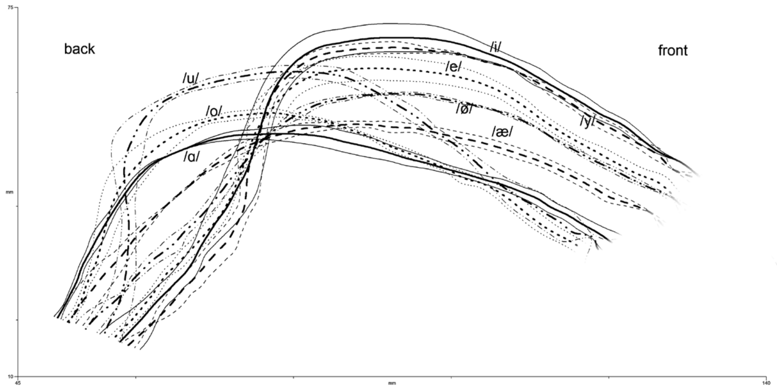

Figure 9. KH102

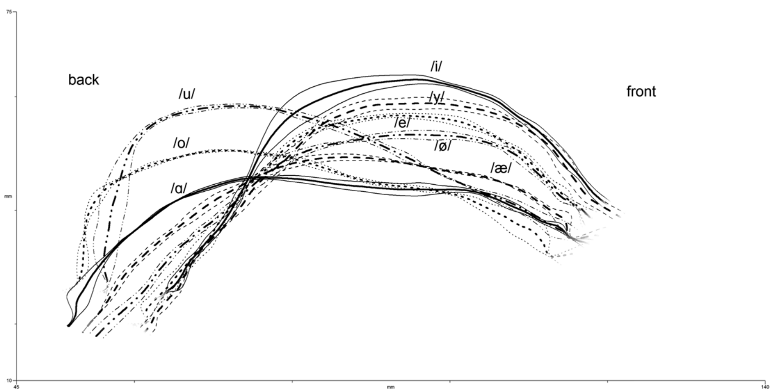

Figure 10. KH108

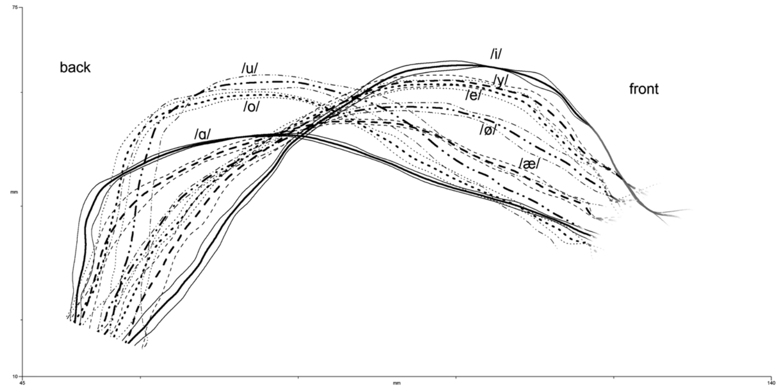

Figure 11. KH112

Figures 9–10. Vowel articulatory systems of three different subjects. Thick lines represent the average articulatory movement of the vowel. Thin lines represent the standard deviation of the vowel.

When observing all vowel articulatory movements together, it seems that back close vowels are more open than front close vowels. In fact, they can not be compared here, because it is not possible to proportion them to the palate and study if they are equally closed or not. The back of the mouth's roof curves downwards, thus narrowing the passage. For further studies, it is important to let the subject swallow some water at the beginning of the experiment to make it possible also to draw a spline to indicate the palate.

If the traditional phonological Finnish vowel diagram (Figure 12) is compared to the vowel systems of these three subjects, there are some points worth noticing. A typical articulatory vowel chart describes symmetric articulation, which is the conventional way to represent the relation of articulatory motor coordination of the tongue in different vowels . This description is based on tactile feedback of the tongue and the relational differences of formants in acoustic studies (Joos 1948: 53–54; Jones 1962: 17). Articulatory descriptions of Finnish vowels are based on acoustic measurements and on these general descriptions of vowel articulation. Even though we are not able to compare the tongue surface spline to palate, it is quite clear that there is much more space for the vowels at the front of the mouth than at the back. In the vowel diagram, the relative difference of /u/ and /o/ is presented as an equal distance as between /i/ and /e/. The data of these subjects shows that the difference between /u/ and /o/ is actually much smaller. Figure 13 presents a vowel diagram, which describes Finnish vowel system as it can be seen based on the data of this study.

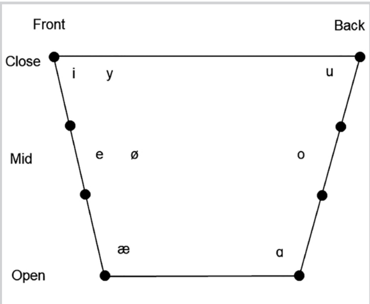

Figure 12. Finnish phonological vowel system presented according to Suomi et al. (2008)

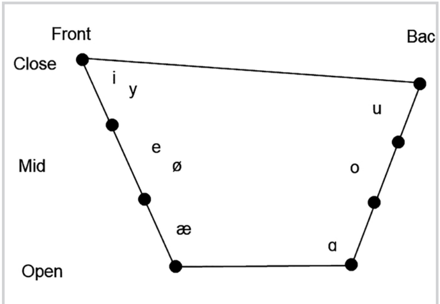

Figure 13. Articulatory vowel chart of. Finnish vowels based on data of this study.

3.3. Rounded and unrounded front vowels

Figures 14–19 contain average tongue surface splines of four vowels from three subjects (KH103, KH106 and KH107). As can be seen from the Figures 14, 15 (KH103) and 18, 19 (KH107) close front vowels /i/ and /y/ and mid-front vowels /e/ and /ø/ differ not only in the rounding of the lips, but also in the articulatory movement of the tongue. It seems that the tip of the tongue is at the same position, but in rounded vowels /y/ and /ø/, the dorsum of the tongue is lowered in comparison to unrounded vowels /i/ and /e/. In Figures 16 and 17 (KH106) this same difference can be found in mid-front vowels, but not in close vowels. This subject produces the difference between /i/ and /y/ only by rounding the lips. This indicates that speakers may use different articulatory gestures for producing the required acoustic characteristics differentiating unrounded and rounded vowels.

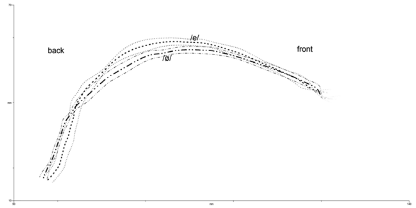

Figure 14. KH103 /e/ (dotted line), /ø/ (dash-dot line)

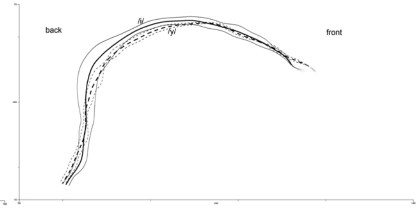

Figure 15. KH103 /i/ (solid line), /y/ (dashed line)

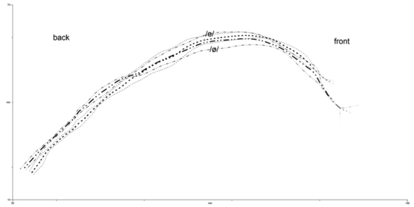

Figure 16. KH106 /e/ (dotted line), /ø/ (dash-dot line)

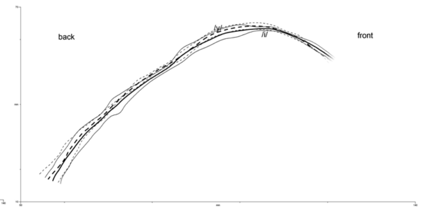

Figure 17. KH106 /i/ (solid line), /y/ (dashed line)

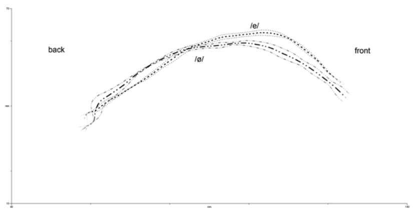

Figure 18. KH107 /e/ (dotted line), /ø/ (dash-dot line)

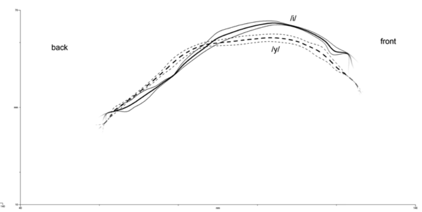

Figure 19. KH107 /i/ (solid line), /y/ (dashed line)

Figures 14–19. Closed and mid-front rounded and unrounded vowels of three different subjects. Thick lines represent the average articulatory movement of the vowel. Thin lines represent the standard deviation of the vowel.

4. Conclusion

The purpose of this study was to compare Finnish vowel articulatory movements to the description of the Finnish vowel system. We found out that the vowel articulation is consistent: the within-subject variation in one vowel is minimal and despite the between-subject variation, the vowel articulatory movements work as a system. The description of vowel articulation has mostly been based on tactile feedback of tongue. On the other hand, the lip articulation is the most visible part of the articulation, and together with tactile feedback of the tongue, it has had an influence on the description of, for example the Finnish front vowels. During talking, it is possible to see the lips and sometimes the tip of tongue, but not the back. We found out that in the Finnish vowel system lip rounding is not always the only movement separating rounded vowels from unrounded ones in close and mid-front vowels. It would be interesting to study the role that lip articulation plays in languages like Swedish and Estonian, which have central rounded vowels. The question is whether acoustic differences in vowels are made with the tongue or with the lips or does it depend on the speaker as it seems to be in Finnish. If so, it seems that an acoustic contrast between two vowels can be produced with different kinds of articulatory movements, despite the fact that the vowel articulatory system is consistent. In addition, it is important to notice that the ultrasound stabilization headset may constrict the movements of the jaw and lips and hence affect to the articulatory movements.

From acoustic studies, it is known that vowel formant frequencies vary depending at context. It suggests that there is some variation in articulatory movements as well. In Finnish, the coarticulatory effect has traditionally been described in a way that the consonant has the coarticulatory effect on vowel. Only by studying articulatory movements it is possible to find out if coarticulation is more like a two-way effect from consonant to vowel and vice versa. Delattre (1965: 88) has found out that the quality of /l/ in American English is dependent on if it is post-vocalic or post-consonantal position. Lin et al. (2012) have studied coarticulatory effect in articulatory movements of English vowels in dental lateral and dental plosive contexts and found differences in some vowels. In this study, we had only one speech sound context to avoid coarticulatory effects and to be able to have the focus only on vowel articulatory movement. We supposed that the dental context we used especially affects to the back vowels by pulling them forward. The Finnish language-specific articulatory vowel chart discussed earlier could be different if vowels were studied in a velar context or without any consonantal context. According to Delattre (1965: 64) coarticulation also affects to the length of the vowel, and as discussed before, it is possible that the length of the vowel affects articulatory movement.

As we have seen, it could be possible to describe vowel articulation in a more specific way according to the shape of the mouth, e.g. in the nature that Joos (1948: 53–54) presented this kind of articulatory-based vowel chart of French. Language-specific vowel charts could help a learner to pronounce sounds of different languages. In that way, it would be easier to compare the vowel systems of the native language and foreign language and so have a better understanding of vowel articulation in the target language.

References

Anderson, Insung, O’Grady, Choo 2004 – Anderson, Victoria; Ko, Insung; O´Grady, William; Choo, Miho. A palatographic investigation of place of articulation in Korean coronal obstruents. *Korean Linguistics* 12, 1–24.

Cleland, Scobbie, Wrench 2015 – Cleland, Joanne; Scobbie, James M.; Wrench, Alan A. Using ultrasound visual biofeedback to treat persistent primary speech sound disorders. *Clinical Linguistics & Phonetics* 29, 8-10, 575–597.

Cleland, Scobbie, Heyde, Roxburgh, Wrench 2016 – Cleland, Joanne; Scobbie, James M.; Heyde, Cornelia J.; Roxburgh, Zoe; Wrench, Alan A. Covert contrast and covert errors in persistent velar fronting. *Clinical Linguistics & Phonetics* 31, 35–55.

Fant 1968 – Fant, Gunnar. Analysis and synthesis of speech processes. In: Bertil Malmberg, ed-in-chief. *Manual of Phonetics*. Amsterdam: North-Holland Publishing Company, 1968, 173–277.

Grönroos 2010 – Grönroos, Antti. *Ultrasonically Enhanced Disintegration. Polymers, Sludge, and Contaminated Soil*. Jyväskylä: University of Jyväskylä, 2010.

Hoole 2016 – Hoole, Philip. *Experimental Studies of Laryngeal Articulation.* München: Fakultät für Sprach- und Literaturwissenschaften der Ludwig-Maximilians-Universität, 2016.

Johansson, Sundberg, Wilbrand 1982 – Johansson, C; Sundberg, J; Wilbrand, H. X-ray study of articulation and formant frequencies in two female singers. *STL-QPSR* 23, 4, 117–134.

Jones 1962 – Jones, Daniel. *An Outline of English Phonetics*. 9th edition. New York: G. E. Stechert & Co, 1962.

Ladefoged 1962 – Ladefoged, Peter. *Elements of Acoustic Phonetics*. London: Oliver and Boyd Ltd, 1962.

Ladefoged 2001 – Ladefoged, Peter. *Vowels and Consonants*. Malden: Blackwell Publishing Ltd, 2001.

Lin, Palethorpe, Cox 2012 – Lin, Susan; Palethorpe, Sallyanne; Cox, Felicity. An ultrasound exploration of Australian English /CVl/ words. In: Cox, Felicity; Demuth, Katherine; Lin, Susan; Miles, Kelly; Palethorpe, Sallyanne; Shaw, Jason; Yuen, Ivan (eds.). *Proceedings of the 14th Australasian International Conference on Speech Science and Technology*. Canberra: Australian Speech Science and Technology Association, 2012, 105–108.

Nakai, Turk, Suomi, Granlund, Ylitalo, Kunnari 2012 – Nakai, Satsuki; Turk, Alice E.; Suomi, Kari; Granlund, Sonia; Ylitalo, Riikka; Kunnari, Sari. Quantity constraints on the temporal implementation of phrasal prosody in Northern Finnish. *Journal of Phonetics* 40, 6, 796–807.

Nakai, Suomi, Wrench 2015 – Nakai, Satsuki; Suomi, Kari; Wrench, Alan. F1/F2 targets for Finnish single vs. double vowels. In: The Scottish Consortium for ICPhS 2015 (ed.). *Proceedings of the 18th International Congress of Phonetic Sciences. August 10th–14th, Glasgow, United Kindom.* Available online: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0471.pdf [accessed on 13.4.2017].

O'Dell 2003 – O'Dell, Michael. *Intrinsic Timing and Quantity in Finnish*. PhD dissertation. Tampere: University of Tampere, 2003.

Proctor, Lo, Narayanan 2015 – Proctor, Michael; Lo, Chi Yhun; Narayanan, Shrikanth. Articulation of English vowels in running speech: A real-time MRI study. In: The Scottish Consortium for ICPhS 2015 (ed.). *Proceedings of the 18th International Congress of Phonetic Sciences. August 10th–14th, Glasgow, United Kindom*. Available online: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2015/Papers/ICPHS0220.pdf [accessed on 13.4.2017].

Scobbie, Wrench, van der Linden 2008 – Scobbie, James M.; Wrench, Alan A.; van der Linden, Marietta L. Head-probe stabilisation in ultrasound tongue imaging using a headset to permit natural head movement. In: Sock, Rudolph; Fuchs, Susanne; Laprie, Yves (eds.). *Proceedings of ISSP2008, December 8th–12th 2008, Strasbourg, France*. Available online: http://issp2008.loria.fr/Proceedings/PDF/issp2008-87.pdf [accessed on 13.4.2017].

Sovijärvi 1967 – Sovijärvi, Antti. *Kielen artikulaatioliikkeistä röntgenkuvien valossa. Helsingin yliopiston fonetiikan laitoksen julkaisuja* 21. Helsinki: Helsingin yliopisto, 1967.

Stevens 1972 – Stevens, Kenneth N. The quantal nature of speech: Evidence from articulatory-acoustic data. In: David, Edwars E. Jr.; Denes, Peter B. (eds.). *Human Communication: A Unified View*. Kingsport: McGraw-Hill Book Company, 1972, 51–66.

Suomi 1990 – Suomi, Kari. *Johdatusta puheen akustiikkaan. Logopedian ja fonetiikan laitoksen julkaisuja* 4. Oulu: Oulun yliopisto, 1990.

Suomi, Toivanen, Ylitalo 2008 – Suomi, Kari; Toivanen Juhani; Ylitalo, Riikka. *Finnish Sound Structure. Phonetics, Phonology, Phonotactics and Prosody*. Oulu: University of Oulu, 2008.

Wood 1982 – Wood, Sidney A J. X-ray and model studies of vowel articulation. *Blom* 23. Lund: Lund University, 1982.

Wrench, Cleland, Scobbie 2011 – Wrench, Alan A; Cleland, Joanne; Scobbie, James M. An ultrasound protocol for comparing tongue contours: Upright vs. supine. In: Lee, Wai-Sum; Zee, Eric (eds.). *Proceedings of the 17th International Congress of Phonetic Sciences, August 17th–21th, Hong Kong, China.* Available online: https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2011/OnlineProceedings/RegularSession/Wrench/Wrench.pdf [accessed on 13.4.2017].