Acoustics'08 Paris
June 29-July 4, 2008
www.acoustics08-paris.org

Magnetic Resonance investigation of palatalized stop consonants and spirants in Russian

G. Ye. Kedrova, N. V. Anisimov, L. M. Zaharov and Y. A. Pirogov

Moscow State Lomonosov University, GSP-2, Leninskije Gory, MGU, 1st Humanities, R. 983, Centre for New Technologies in Humanities, 119992 Moscow, Russian Federation
kedr@philol.msu.ru

The research in question is aimed at the experimental MRI-investigation of articulatory gestures corresponding to the soft Russian stop consonants [p’], [t’], [k’], [b’], [d’], [g’] and spirants [s’], [z’], [h’] countered with their hard counterparts [p], [t], [k], [b], [d], [g], [s], [z], [h] in Russian speech production.
An experimental dataset was based upon 2D MR-images, audio- and video recordings taken from 4 native speakers of Russian producing VCC**V** sequences (Russian pseudo-words with the second vowel stressed) containing Russian consonants under investigation in the vocalic contexts [a]_[a], [a]_[e], and [i]_[i]. All speaking subjects had standard pronunciation and were without any perceptible articulation disease. MRI investigation of the Russian consonantal phonemes was based upon admitted procedures and techniques though expanded with several new original methods elaborated by the Russian team of experimentalists. Experimental data was collected through several MRI sessions (done in a month and in a year with participation of the same speaking subjects), ensuring thereby credibility and robustness of the experimental results.
A detailed analysis of the whole dataset of MR-images of palatalized Russian consonants countered with their non-palatalized counterparts revealed special articulatory pattern of palatalization in Russian.

1 Introduction

At present, the need for authentic 2- and 3D articulatory images peculiar to various languages is determined by further development of speech production models - especially by an ambitious task of constructing faithful 3D articulatory speech synthesis. The same kind of data is required to respond to the most demanding challenges of innovative CALL applications as well as to other open questions of speech rehabilitation and correction activities. Meeting this needs a MRI investigation of the model articulatory patterns in various languages has been developing at a quick pace and has written since its early stage a long and rather successful history.

Early research in the field dates from the mid-60s of the previous century. From that time onwards a rich pool of empirical data on different types of vocals and consonants (primarily fricatives, approximants, laterals, retroflex and nasals) in French (P. Badin), Swedish (O. Engwall), German (P. Hoole, B. Kröger), British and American variants of the English language (M. Tiede, S. Narayanan) as well as in some non-Indo-European languages (K. Honda, S. Maeda) is available in the literature. Recently another ambitious target - 3D modeling of the speech articulation in dynamics (articulatory synthesis) has determined the main trend in the area. A series of experimental research conducted in the same strain were not long in coming, resulting thus in a very realistic 3D model of speech articulation processes – a 3D articulatory synthesis called “Virtual Talking Head” [1].

However, unlike the latest global trends in experimental and applied phonetics, the MRI investigation of speech production mechanisms, articulatory patterns and dynamics included, is still at its outset in Russia. There is a considerable deficiency of authentic experimental research and/or reliable data for Russian articulatory and co-articulatory patterns either for the vocal or consonant system of the Russian language. Only a few pilot investigations based mainly upon 2D on-line MR-imaging techniques are available in the literature. No detailed data on the nature of Russian palatalization phenomenon that would be extracted from MRI investigation still exist.

However, contrary to some generally accepted opinion based on assertion that palatalization of Russian consonants is a concomitant articulation gesture, antecedent Russian tradition of instrumental experimental research of the palatalization highlighted soft and hard consonants’ articulation patterns as two completely different physiological motor stereotypes. In the 1960s-1970s, L. Skalozub, being the first Russian experimentalist in this field, completed a series of wide-ranging studies of articulation dynamic models using x-ray processing of natural speech. One of the main targets of her large-scale x-ray investigation of Russian consonantal articulation patterns was a comparison of manually traced profiles of the articulatory tract observed during phonation of various consonants of the Russian language, soft and hard consonantal counterparts included [2]. Thus, she stated that the vocal tract contours drawn upon x-ray filming of soft and hard consonants in experimental sessions of producing natural and pseudo speech stimuli differed drastically either in stationary phase of a sound or in its dynamic deployment (motor stereotype). Furthermore, these vocal tract configurations revealed a sort of commonalities in the overall contour typical for all palatalized consonants, that L. Skalozub defined as dorsal articulation mode coupled with pre-palatal location [3].

Substantial insufficiency in this data prevented Skalozub from more specific and definite conclusions, though the idea itself was estimated as very promising. Since x-ray investigation is an empirical method that is very harmful and extremely dangerous to the human health, many research topics of L. Skalozub and her disciples are, until now, in an incomplete state, therefore numerous significant aspects of the Russian speech articulation remain still unexplored. One of the urgent topics deals still with ambivalent nature of the Russian palatalization. It should be also mentioned that until now there is no 2D or 3D MRI techniques applications targeted to the studies of the Russian soft and hard consonants’ articulation. Therefore our current research aims to fill this gap.

2 Material and Methods

The research in question forms an integral part of a broader magnetic resonance imaging investigation of the complete inventory of articulatory motor patterns representative for contemporary Russian language pronunciation practice. According to our main targets special interest was focused upon the Russian palatalization phenomenon, since linguistic theory as well as long-term practice of linguistic applications proved Russian soft consonants to be one of the most unique characteristic features of the language system and definitely the most difficult pronunciation skill to acquire by a foreigner.

Up to now our investigation of the Russian soft and hard articulatory patterns was grounded on 2D MRI technology (though we are planning to enlarge this methodology to acquire 3D MRI in the future). MR images' database was assembled from 2D on-line MR-images of the sagittal cut of articulatory tract enlarged with audio- and video recordings taken from native speakers of Russian with standard pronunciation and without any perceptible articulation disease. The MRI investigation of the Russian sound patterns (phonemes) was based upon admitted procedures and techniques which were expanded with several new original methods offered and successfully tested by the Russian team of experimentalists [4]. This primary knowledge was used as a keystone for MRI experiments on soft and hard consonantal articulation patterns' confrontation, as well as for further processing of the collected data.

According to previously approved experimental instructions each speaking subject was asked to produce a series of VCC**V** sequences (Russian pseudo-words with the second vowel stressed) containing all Russian consonants in the [a]_[a], [i]_[i], [a]_[e] vocalic frameset, that is a[b:]**a**, a[d:]**a**, etc.; a[b':]**a**, a[d':]**a**, etc.; i[b':]**i**, i[d':]**i**, etc.; a[bb]**e**, a[dd]**e**, etc, a[b':]**e**, a[d':]**e**, etc, repeating each stimulus several times during a session of MR image acquisition. Our main interest was focused on the [a]_[a] and [a]_[e] contexts as minimal distinctive positions for hard and soft consonantal correlates. In this paper we discuss articulatory patterns drawn upon corresponding MR images typical for soft Russian stop consonants [p'], [t'], [k'], [b'], [d'], [g'] and spirants [s'], [z'], [h'] confronted with their hard counterparts [p], [t], [k], [b], [d], [g], [s], [z], [h].

2.1 MRI specifications

Articulation patterns of hard and soft Russian consonantal correlates were investigated after having been traced upon articulation data collected from 4 native speakers of Russian: 3 males (LZ, AE, ON) and a female (GK) using Magnetic Resonance Imaging (MRI). MRI experiments were realized in Educational-Research Center of Magnetic Tomography and Spectroscopy of the Moscow State Lomonosov University on a 1.5 T MR system (Tomikon S50 «Bruker»). The receiver coil was a quadrature neck coil.

All speaking subjects were lying in supine position, any special mechanism of the head fixation not provided. They were required to repeat experimental stimuli at their own pace as many times as possible during acquisitions of MR images. The experimenter instructed the subject to initiate the speech process by counting every second, a couple seconds before the MRI acquisition starts.

Consistently with generally accepted dynamic MRI techniques, we've arranged simultaneous audio recordings taken via a microphone LifeVideo(tm) fixed on a receiver's coil close to the speaker's mouth. As this recording was strongly dominated by the MR scanning machine noise, it was impossible to label reliably and in details the speech sequences, therefore a parallel recording of the starting points of MRI sequences was also previewed. Both recordings were presented as a two-channel oscillogram (Fig. 1), which enabled more precise timing of an MR image with a particular phase of phonation, as well as with any phase of pausal period in the future procedures of images' identification. Other relevant details of image acquisition and image analysis are similar to those characterized in details by Kedrova, et al. in [5]. Irrespective of the MRI sessions extra control audio and video recordings of the same speech data from the same reference subject producing speech in the same position were realized in a professional record's studio environment.

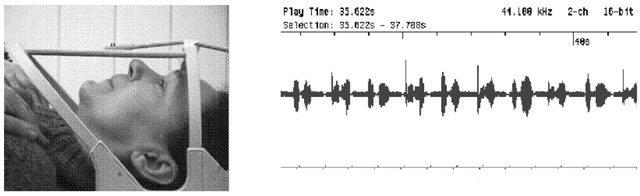

Fig.1 Position of a reference subject GK in MRI experiment. To the right: two-channel oscillogram displaying phonation and starting points of MRI sequences.

The total data set was collected in two separate experimental sessions with the time gap of one month and one year, but using the same reference subjects and with the same language stimuli for all acquisitions.

In our experiments on MRI research of articulatory models of Russian hard and soft consonants, MR scanning was executed on sagittal slice with the slice thickness of 9 mm and to a field of view 200x120 mm. Well known pulse sequence "gradient echo" was used with the following parameters: TR=12 ms, TE=5.5 ms, FA=10 degrees. Under these conditions it was possible to obtain MR images with 2-2,7 frames in a second and with 3 mm in-plane resolution.

The whole data set of MR images collected in all experimental MRI acquisitions was identified and ascribed to each phase of a phoneme realization. It is worth mentioning that in all experimental sessions we've observed the highest degree of image matching within each speaker's various performances of a particular consonantal phoneme under investigation, though certain dependencies from a vocal context were also observed. Thus, we consider our results as another proof of the linguistic concept of phoneme defined as a "psychomotor complex formed in the early childhood via association of contiguity" [6].

2.2 Russian hard and soft consonantal phonemes' articulation contours

It has already become common knowledge that human motor system offers an extremely rich set of possibilities for executing a given task, therefore any stable motor stereotype should be considered as an evidence of a special target-oriented low-cost motor pattern. The observed in the experimental data motor stereotyping presumably resulted from a phoneme's inherent properties and far less – from a specific phonetic context. Moreover the experimental data presented a considerable grade of articulatory contours' matching for each type of voiced / unvoiced hard and soft consonant irrespective of most vocalic contexts (the only exception worth mentioning has been detected for certain consonantal phonemes in the front-vowels frameset). Typical vocal tract configuration patterns of voiced and unvoiced hard and soft consonants under investigation are exposed on figure 2. MR images on the figure represent pronunciation mode of one of the experimental subjects (subject LZ, a male from Moscow origin).

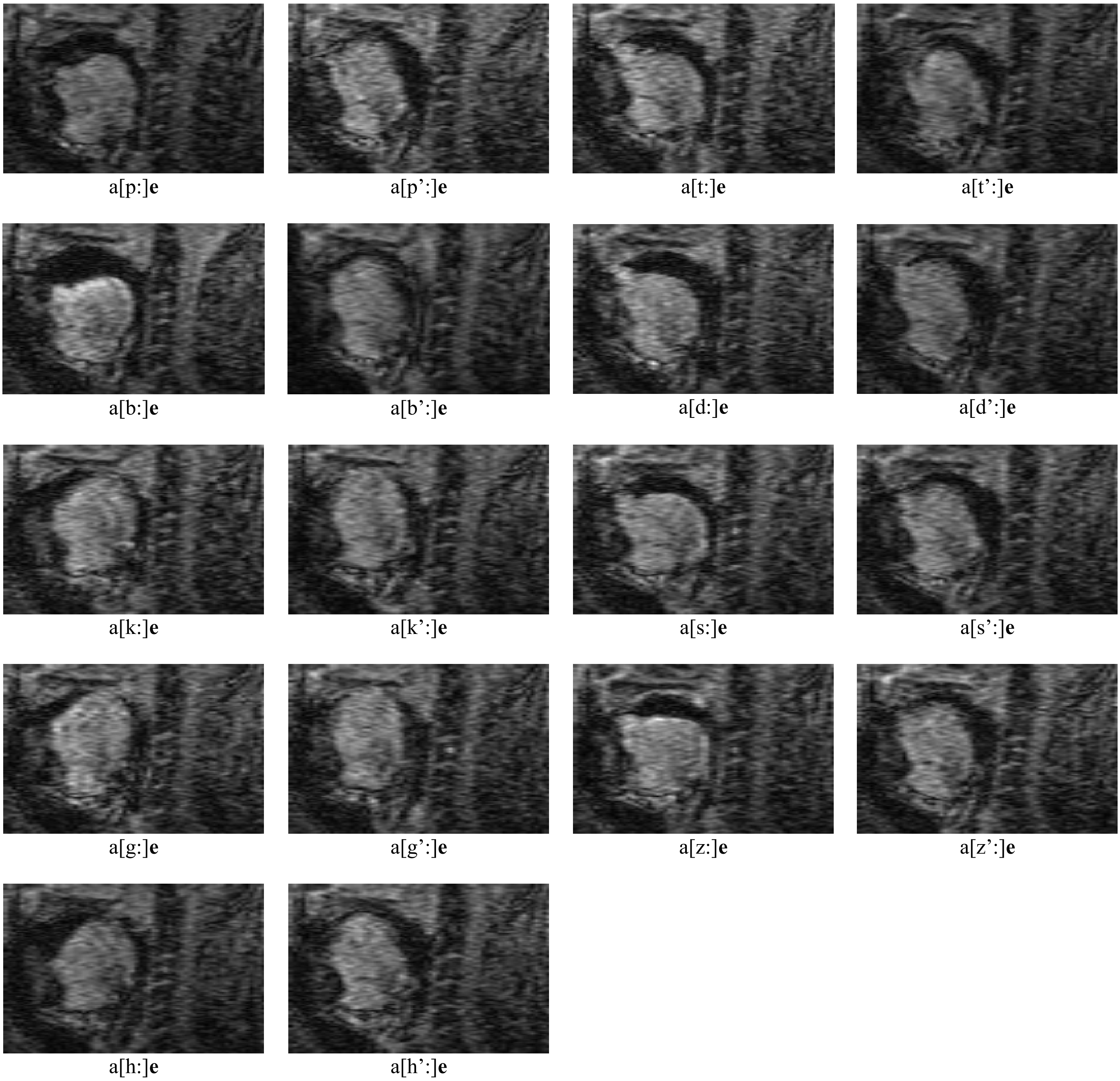

Fig.2 MR images of the stationary phase of the consonants [p], [p’], [b], [b’], [t], [t’], [d], [d’], [k], [k’], [g], [g’], [s], [s’], [z], [z’], [h], [h’] in the [a]_[**e**] frameset (speaking subject LZ).

Another nameworthy observation upon pronunciation mode of the subject LZ deals with apparent less diversity of the tongue shape observed in soft consonants compared with their hard counterparts. While position and form of the tongue shape and body characteristic for every hard consonant highly depend on its role in the entire sound system of the language associated with a certain bundle of distinctive features, all soft consonants seem to be produced with a roughly similar articulatory pattern characterized by high degree of tongue body constriction coupled with the tongue dorsum huddled up in the front region of the oral cavity. The raising gesture of the tongue dorsum entails also a considerable movement forward of the tongue root expanding thus very significantly the pharyngeal cavity.

Therefore, we could state that there are several distinct differences in articulation configurations between hard and soft consonants in Russian observed in all consonantal articulations on our MR images. These differences appear to be more prominent for coronal plosives [t] and [d] while the articulatory contour reveals high degree of involvement in sound production of the tongue tip as well (and very anterior part of the blade), forming a closure in the alveolar region.

Similar observation should be made upon examination of articulatory patterns of fricatives [s] and [z], though their articulation could be also described as significantly more velar.

Russian hard bi-labial stops [b] and [p] are also distinctly opposed to their soft counterparts with the displacement of the main tongue body to the posterior region of the articulatory tract, alongside with the characteristic hollow on the tongue. This hollowing necessarily makes the overall tongue shape concave. It is also worth to remind that these consonants are in Russian phonetic tradition defined as labiovelar phonemes.

At the next stage of our research we have compared atriculatory patterns of hard and soft Russian consonants under current investigation in all the four experimental speaking subjects. The data collected in all series of experiments evidence for existence of at least two different articulatory-motor models of Russian palatalization phenomenon. Most essential difference was observed in pronunciation patterns of almost all the soft and several hard consonants produced by a female subject born and raised up outside the Moscow region (in a city located in the South-Eastern part of Russia). Typical vocal tract configuration patterns of voiced and unvoiced hard and soft consonants under investigation realized by this subject (subject GK) are exposed on figure 3.

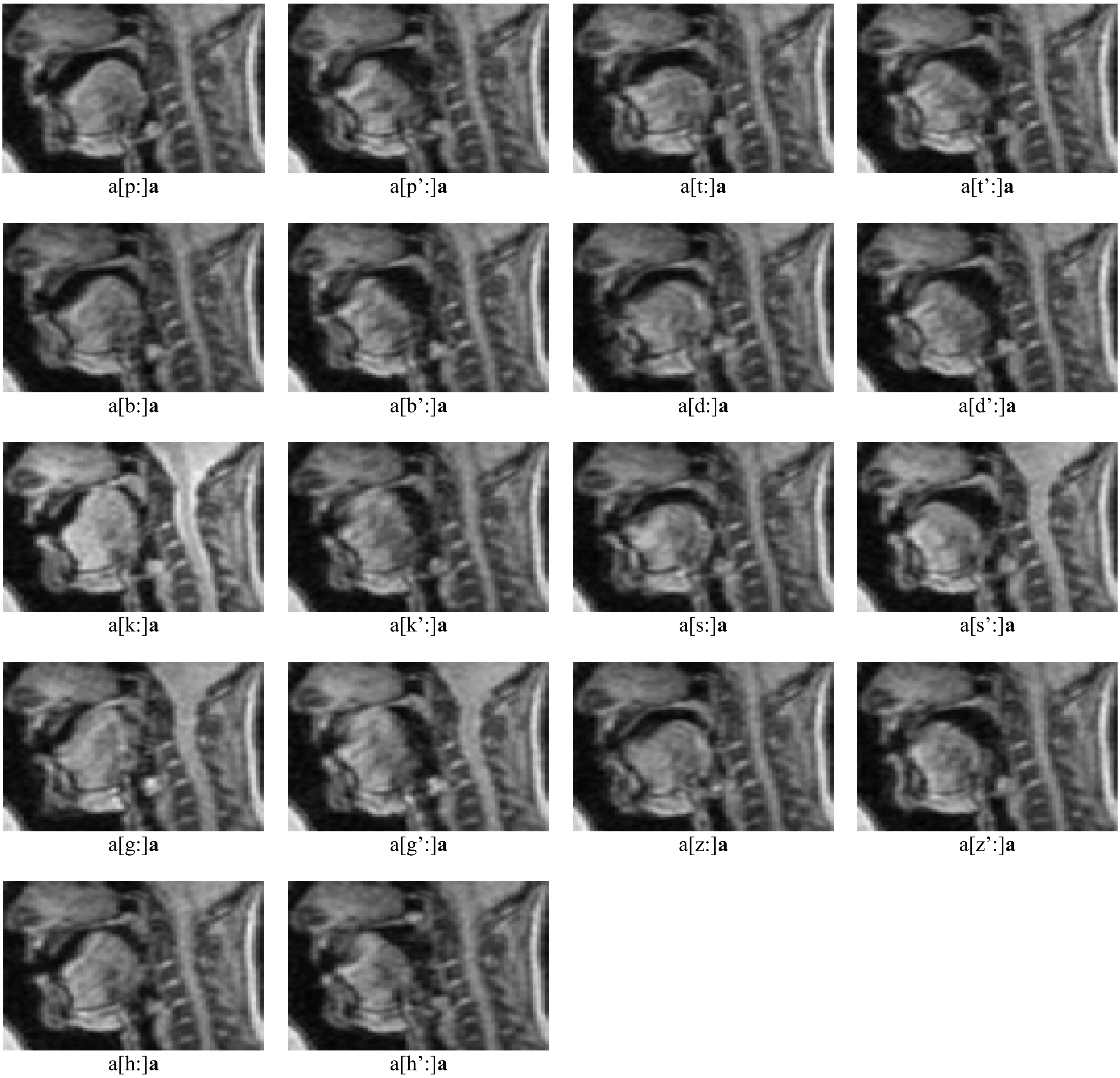

Fig.3 MR images of the stationary phase of the consonants [p], [p’], [b], [b’], [t], [t’], [d], [d’], [k], [k’], [g], [g’], [s], [s’], [z], [z’], [h], [h’] in the [a]_[**a**] frameset (speaking subject GK).

3 Discussion

Comparative study of speaking subjects' consonantal articulation contours reveals certain similarities in the production of hard and soft consonants as well, especially for voiced and unvoiced paired phonemes. It is obvious, that similar to the subject LZ articulation mode voiced and unvoiced consonantal pairs demonstrate quasi-identical articulatory configurations irregardless of them being either hard or soft.

However, contrary to the previously described articulation patterns of palatalization, soft consonants' articulation in the speech production of subject GK demonstrates more robust diversification in articulatory-motor stereotypes both in general or for particular types of sound. It concerns primarily articulation patterns of the soft counterparts of back consonants [k] and [g], as well as soft counterparts of the spirants [s] and [z]. These spirants demonstrate distinct type of tongue body contour which distinguishes them from other soft phonemes in the pronunciation of this speaking subject.

In general, one could state that the basic tongue configuration for palatalized phonemes in the subject GK phonation is essentially less "monolithic" and apparently less tight, while the most active part of tongue articulation is focused on the tongue blade and surrounding tissue. This aspect favours tongue tip's involvement in various types of articulatory activities, as it has been earlier stated that the degree of the dorsum's involvement in articulation activity effectively limits the freedom of the tongue tip and/or blade activities [7].

4 Conclusion

The methodological and technological approach elaborated in current investigation has proven its validity and could be recommended for implementation. Preliminary results of our research support the definition of palatalization in Russian as a particular integral articulatory motor mechanism. A detailed analysis of the whole dataset of MR-images of several palatalized Russian consonants countered with their non-palatalized counterparts in pronunciation practice of Muscovites and a speaker of non-Moscow origin revealed two basic models of articulatory activity during phonation of palatalised consonants.

One model should be characterized as based upon massive involvement of the tongue body in all articulation processes coupled with substantial raising of the tongue dorsum in the front region of the oral cavity and restriction of the tongue root, while the other one seems to be focused mainly on the tongue blade activity that is locally realised in its convex shape in alveolar and post-alveolar regions.

We would suggest that the described differences could become one the most influential factors in development of the database containing various motor stereotypes of coarticulation patterns observed in fluent Russian speech, either in certain consonantal clusters or within the CV collocations.

Further investigation of main coarticulative effects dependant on various models of palatalization described in our paper might be of great interest for the Russian dialectal phonetics as well.

The MRI data collected through the current research would make a significant contribution for further development of 2D and 3D modeling of the Russian articulation based on MRI data, as well as elaboration of efficient computer simulation technology for visualization of human articulatory strategies. Thus constructed animations should be recommended for improvement of methodology and practice in teaching and learning foreign languages (in our case, the Russian language). Our data could be also used in those speech synthesis systems that pretend to be based on physiologically relevant articulatory models.

Acknowledgments

The research reported here and the preparation of the manuscript and poster presentation were partly supported by NWO-RFBR grant No. 047.011.2005.017 (Dutch-Russian cooperation program). We thank all our subjects for participation. The authors are very grateful to the colleagues from Philological Faculty and Faculty of Physics of Moscow State Lomonosov University, and to colleagues from the Institute of Phonetic Sciences of the University of Amsterdam (Dr. Prof. Cecilia Ode in particular) for collaborative and fruitful discussions.

References

[1] P. Badin, A. Serrurier, Three-dimensional modeling of speech organs: Articulatory data and models, *Transactions on Technical Committee of Psychological and Physiological Acoustics, Acoustical Society of Japan, Kanazawa, Japan, 36*(5), 421-426 (2006)

[2] Л.Г. Скалозуб, Динамика звукообразования (по данным кинорентгенографирования), Киев (1979):

[3] Л.Г. Скалозуб, Палатограммы и рентгенограммы согласных фонем русского литературного языка, Киев (1963)

[4] Н.В. Анисимов, Г.Е. Кедрова, Л.М. Захаров, Ю.А. Пирогов, МРТ-визуализация процессов артикуляции при порождении речи, *II Евразийский конгресс по медицинской физике и инженерии, Медицинская физика – 2005,* Москва, 239-240 (2005)

[5] G. Kedrova, N. Anisimov, L. Zaharov, Ju. Pirogov, Contrastive study of the MRI representation of Russian vowel articulation (against French, German and Korean analogues), *Proceedings of the 33rd International Acoustical Conference*, 105-108, (2006)

[6] В.А. Богородицкий, Очерки по языковедению и русскому языку, Едиториал УРСС: М., (2004)

[7] M. Nakamura, Articulatory constraints and coronal stops: an EPG study, *INTERSPEECH-2005*, 1017-1020 (2005)