An approach to real-time magnetic resonance imaging for speech production[a]

Shrikanth Narayanan
Departments of Electrical Engineering, Computer Science, and Linguistics, University of Southern California, Los Angeles, California 90089

Krishna Nayak[b]
Department of Electrical Engineering, Stanford University, Stanford, California 94305

Sungbok Lee
Departments of Electrical Engineering and Linguistics, University of Southern California, Los Angeles, California 90089

Abhinav Sethy
Department of Electrical Engineering, University of Southern California, Los Angeles, California 90089

Dani Byrd[c]
Department of Linguistics, University of Southern California, Los Angeles, California 90089

(Received 29 July 2003; revised 2 January 2004; accepted 13 January 2004)

Magnetic resonance imaging (MRI) has served as a valuable tool for studying static postures in speech production. Now, recent improvements in temporal resolution are making it possible to examine the dynamics of vocal-tract shaping during fluent speech using MRI. The present study uses spiral k-space acquisitions with a low flip-angle gradient echo pulse sequence on a conventional GE Signa 1.5-T CV/i scanner. This strategy allows for acquisition rates of 8–9 images per second and reconstruction rates of 20–24 images per second, making veridical movies of speech production now possible. Segmental durations, positions, and interarticulator timing can all be quantitatively evaluated. Data show clear real-time movements of the lips, tongue, and velum. Sample movies and data analysis strategies are presented.
[DOI: 10.1121/1.1652588]

PACS numbers: 43.70.Jt [AL] Pages: 1771–1776

I. INTRODUCTION

A perennial challenge in speech production research is the ability to examine real-time changes in the shaping of the vocal tract. These deformations in the vocal tract serve two critical functions in speech production: (1) Creation of supraglottal sources through the coordination of appropriate articulator configurations and aerodynamic conditions, and (2) acoustic filtering of the laryngeal and supraglottal sources by actively modifying the shape of the vocal tract to produce distinctive spectral patterns for the different speech sounds. Hence, spatiotemporal information about speech movements is critical not only to understanding and modeling the speech production process but also to a thorough understanding of speech acoustics. Many approaches are available to the researcher for obtaining information about the rapid and complex movements of the mouth and face that participate in creating speech sounds—e.g., electropalatography (EPG) to examine linguapalatal contact, point-movement tracking (e.g., x-ray microbeam, magnetometry) for dynamic oral information, or ultrasound for examining tongue-surface contours in the mouth and pharynx. However, none of these techniques yields real-time moving images of articulators along the entire length and diameter of the vocal tract.[1] We report for the first time a novel high speed MRI technique for imaging the moving vocal tract in real time.

A. Prior work on dynamic imaging

Cine x-ray techniques had been popular in speech research to obtain lateral (midsagittal) images of the vocal tract (Fant, 1960; Delattre and Freeman, 1968; Perkell, 1969; Subtelny *et al.*, 1972; Giles and Moll, 1975). Cross-sectional vocal-tract areas (area functions) were estimated from these midsagittal images for acoustic modeling. The use of such radiographic techniques has declined significantly, mainly due to the radiation risks involved. Computer-aided tomography (Kiritani *et al.*, 1977) is capable of yielding cross-sectional information but still suffers from radiation risks and relatively low speed of imaging. Ultrasound and MRI have proved to be viable alternatives for such investigations.

Ultrasound provides an acceptable means of studying tongue shapes and movements during speech production (Stone, 1990). The technology is safe and noninvasive, and the imaging speed is suitable for studying the dynamics of speech production. However, the entire vocal tract cannot be studied at once by this method. Moreover, due to the presence of the airway above the tongue, the palate cannot be imaged, while due to the presence of air space below, the tongue tip/blade cannot be successfully captured. Neverthe-

---

[a] Portions of this work were presented at the 2003 Nashville meeting of the Acoustical Society of America [Narayanan *et al.*, J. Acoust. Soc. Am. **113**(4,2), 2258 (2003)].
[b] Now affiliated with the USC Department of Electrical Engineering.
[c] Electronic mail: dbyrd@usc.edu.

less, modified ultrasound techniques have helped to further the understanding of the 3D model of tongue, either by using multiple scanning procedures (Watkin and Rubin, 1989) or by using other parallel instrumentation measurements such as x-ray microbeam (Stone, 1990) or palatography (Stone *et al.*, 1992).

Magnetic resonance imaging (MRI) is a powerful tool for obtaining vocal-tract geometry data and does not involve any known radiation risks. The images have good signal-to-noise ratio, are amenable to computerized 3D modeling, and provide excellent structural differentiation. In addition, the tract (airway) area and volume can be directly calculated. The low image-sampling rate, however, has largely limited MRI to the study of sustained speech sounds, corresponding to ''static'' tract shape (Baer *et al.*, 1991; Greenwood *et al.*, 1992; Moore, 1992; Sulter *et al.*, 1992; Dang *et al.*, 1993; Narayanan *et al.*, 1995; Story, 1995; Narayanan *et al.*, 1997; Alwan *et al.*, 1997; Ong and Stone, 1998, Narayanan *et al.*, 1999). Some of these studies used multiple repetitions of an utterance to reconstruct the 3D volume of the vocal tract images from static postures. These data have been valuable in providing hitherto unknown details of the 3D vocal-tract morphology and interspeaker variation during speech production. MRI data have also been used to improve models of the acoustics of speech sounds including vowels (Story *et al.*, 1996), liquids (Bangayan *et al.*, 1996; Espy-Wilson *et al.*, 2000), and fricatives (Narayanan and Alwan, 2000).

In the past few years, improvements in temporal resolution have allowed MRI to move from being limited to imaging static postures (Mády *et al.*, 2001, 2002; Demolin *et al.*, 2000). Progress toward increased temporal resolution in MR imaging is challenged by issues related to poor signal-to-noise ratio (SNR) and susceptibility artifacts when fast imaging protocols, such as fast gradient echo techniques and echo planar imaging, are involved. At this point, it is worthwhile to clarify the use of the terms *dynamic* MRI and *real-time* MRI in the present article. We use the term *dynamic* to refer to the creation of images from an actively articulating subject, rather than a static postural source. We reserve the term *real-time* MRI to refer specifically to directly capturing or acquiring moving image data in real time. That is, *dynamic* refers to the source, *real time* to its acquisition.

The dynamic MRI technique (Foldvik *et al.*, 1993) provides a way to capture valuable kinematic information. It relies on gated scanning on numerous repetitions of the same speech sequence to reconstruct the impression of articulatory movement in time. Note that the reconstructed sequence is drawn from across several repetitions. Using a multiplanar dynamic MRI technique, Shadle *et al.* (1999) formed pseudo-time-varying images of the vocal tract using a simultaneously recorded audio signal and the scanned images acquired while an utterance is repeated. The technique showed positive results in terms of getting fairly accurate volumetric measurements of area functions and tongue volumes.

In the challenging area of real-time MRI in which motion is imaged directly, Demolin *et al.* (2000) have shown significant improvement towards real-time MRI for speech production by the use of an ultrafast implementation of turbo spin echo sequence (TSE) with an acquisition speed of about 4 images/second for a single plane. This initial work on real-time MRI and work by Mády *et al.* (2002) was promising and worth further investigation.

While these MRI advances represent a significant improvement in the quality of information that could be attainted about changes in tongue, lip, and velum positioning over time, they are still not close to the temporal resolution necessary for capturing the dynamic characteristics of tongue movement, which demands a *minimal* sampling rate on the order of 20 Hz. Such high MR imaging rates present significant technical challenges. We present one approach to dynamic real-time MR imaging that successfully addresses these challenges.

II. TECHNICAL DESCRIPTION

Real-time MRI has been developed for several applications including cardiac imaging, abdominal imaging, and interventional imaging, which also demands a high temporal resolution for tracking a moving structure (Kerr *et al.*, 1995; 1997; Santos *et al.*, 2002; Nayak *et al.*, 2001). We pursue real-time imaging of the upper airway, which for speech also requires high spatial and temporal resolution and resilience to image artifacts. We utilized gradient echo imaging with a fast interleaved spiral acquisition strategy. On a conventional 1.5-T imager with high-speed gradients, we were able to acquire images with 110-ms temporal resolution. Critically, this allows effective reconstruction rates of 24 frames per second using a sliding window technique.

A. Experimental methods

Experiments were performed on a GE Signa 1.5-T system with gradients supporting 40-mT/m amplitude and 150-T/m/s slew rate and receiver capable of 4-μs sampling ($\pm$125-kHz receiver bandwidth). A generic head coil, not specially adapted for vocal-tract imaging, was used in all studies. The institutional review board of Standford University approved the imaging protocols, and each subject was screened for magnetic resonance imaging risk factors and provided informed consent in accordance with institutional policy.

Sequences were implemented within a custom real-time imaging (RTI) framework previously described (Kerr *et al.*, 1995; Santos *et al.*, 2002). Interactive control, continuous reconstruction, and display of images were performed on a workstation adjacent to the scanner.

B. Pulse sequence design

Previous approaches to imaging the upper airway in the vocal tract have been limited to standard 2DFT (two-dimensional Fourier transform) and 3DFT sequences (Shellock *et al.*, 1992). These techniques produce good quality static images but are not efficient enough to capture dynamic events. For rapid imaging of the upper airway (UA), we focused on using the spiral readout scheme for accelerating acquisition. Spirals are an alternative scheme for sampling the spatial frequency domain (k-space), in which data are acquired in a spiraling pattern. Twenty interleaved spirals together form a single image. Spirals are highly time effi-

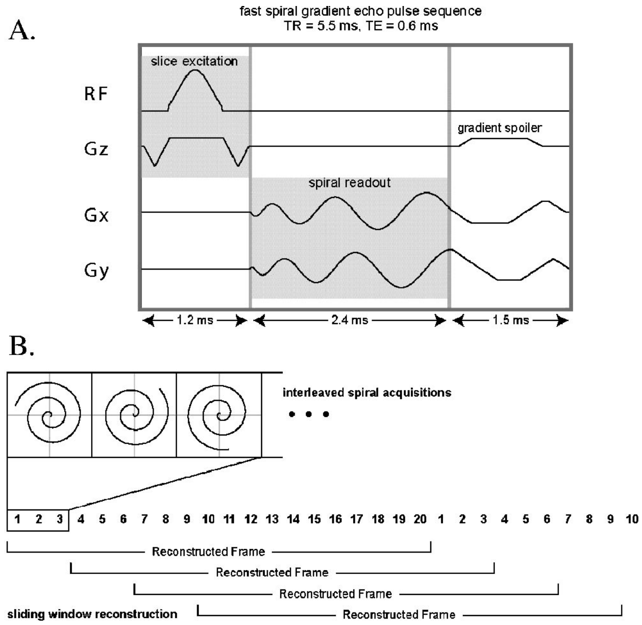

FIG. 1. (A) Pulse sequence used in real-time upper airway imaging. A 1.4-ms slice selective excitation (640-μs rf) is followed by a 2.4-ms spiral readout, and gradient refocusing in X and Y, and crushing in Z. The TR is 5.5 ms. Twenty interleaves are used to achieve 112 pixels over a nominal 20-cm FOV (30-cm FOV was used in most studies). Complete images are acquired every 110 ms, and were reconstructed at 24 frames/s using a sliding window. (B) Acquisition timing. In spiral imaging, the frequency domain is sampled using spiral-shaped trajectories (top). Twenty interleaved spirals are required to form a single image. In dynamic real-time imaging, we continuously reacquire these 20 interleaves (middle). Each image is based on 110 ms of acquired data; however, using sliding window reconstruction, images can be reconstructed at a higher rate as new interleaves become available (bottom). Note that motion within the 110-ms window will result in motion artifacts and/or blurring.

cient (Meyer *et al.*, 1992; Nayak *et al.*, 2001) and have excellent motion properties (Nishimura *et al.*, 1995), as compared to the widely used 2DFT Cartesian approach. Their main limitation is blurring due to off-resonance. This is a particular problem in the UA because the difference in magnetic susceptibility between air and tissue (Schenk, 1996) causes large amounts of off-resonance near to air–tissue interfaces. The amount of blurring experienced in spiral imaging is proportional to the amount of phase that accrues during each readout. To mitigate this effect, we used extremely short 2.4-ms readouts.

The pulse sequence is shown in Fig. 1(A). A 640-μs excitation is followed by 2.4-ms readouts and a gradient spoiler. The total TR is 5.5 ms. Twenty interleaves are required to achieve 1.8-mm resolution over a nominal 20-cm FOV (field of view). A slice thickness of 5 mm was used. (It should be noted that a 3-mm slice thickness would be possible with the same pulse sequence but with a sacrifice in SNR.) In our studies, a 30-cm FOV was used to prevent aliasing artifacts from the large volume captured by the head coil. This resulted in an effective in-plane resolution of 2.7 mm. Complete images were acquired every 110 ms, and were reconstructed at 24 frames per second using a sliding window (Holsinger *et al.*, 1990). A schematic representation of this procedure is given in Fig. 1(B). Basically, this means that while each image is acquired over 110 ms, since data for each image are acquired in pieces, images can be reconstructed more often as portions of the frequency-domain data are updated.

FIG. 3. The [l] constriction in ‘‘Say peal leap again;’’ the real-time MRI movie can be downloaded at sail.usc.edu/production/rtmri/jasa2004.

III. EXAMPLE STUDIES

Two subjects (SN and KN) were recorded in a preliminary study producing normal speech using this protocol. The pulse sequence used is that given in Fig. 1. Study 1, in which both subjects participated, included English sentences varying the syllable position of /n, r, l/. Study 2, in which only SN participated, included sentences in Tamil varying among five liquids. This dataset matches that of a point-movement tracking study found in Narayanan *et al.* (1999).

Results show clear imaging of the entire vocal tract and real-time movements of the lips, tongue, and velum. Seg-

FIG. 2. The [l] constriction in ‘‘Say pea leap again;’’ the real-time MRI movie can be downloaded at sail.usc.edu/production/rtmri/jasa2004.

FIG. 4. The [n] constriction in ‘‘Say bean knee again;’’ the real-time MRI movie can be downloaded at sail.usc.edu/production/rtmri/jasa2004.

FIG. 5. The Tamil retroflex [ɭ] constriction from "...paɭam;" the real-time MRI movie can be downloaded at sail.usc.edu/production/rtmri/jasa2004.

mental durations, positions, and interarticulator timing can all be quantitatively evaluated. Sample movies and data analysis strategies are presented below.

A. English sonorants

The sentences used in Study 1 include
"Say——again."
pea leap, peal heap, peal leap, peal-y, pea reap, pier heap, pier reap, peer-y, be knee, bean he, bean knee, bean-y

Captured frames from movies of the moving vocal tract for three sentences are presented (see Figs. 2–4) and files of the entire real-time movie for each can be downloaded at sail.usc.edu/production/rtmri/jasa2004.

B. Tamil liquids

The sentences used in Study 2 to examine liquid consonants in Tamil embed stimuli of the sort "pali" [sacrifice], "paɭi" [nonce], "paɻi" [blame], "pari" [horse], and "par̠i" [pluck] in a frame utterance; the full set of sentences and a phonetic description can be found in Narayanan *et al.* (1999). Captured frames from movies of the moving vocal tract for the first three consonants given above are presented in Figs. 5–7, and files of the entire real-time movie for each can be downloaded at sail.usc.edu/production/rtmri/jasa2004.

It is noteworthy that Movie 5 indicates the temporal as well as spatial extent of the retroflex constriction and release. In Fig. 8, a selection from this movie is presented to allow for the inspection of the vocal-tract shape changes over multiple frames simultaneously. It indicates that the tongue tip curling for the retroflex lateral in [paɭam] begins well before the *preceding* labial closure is achieved.

FIG. 6. The Tamil [l] constriction from "... palam;" the real-time MRI movie can be downloaded at sail.usc.edu/production/rtmri/jasa2004.

FIG. 7. The Tamil [ɻ] constriction from "...paɻam;" the real-time MRI movie can be downloaded at sail.usc.edu/production/rtmri/jasa2004.

IV. SKETCH OF ANALYSIS CHALLENGES

The method presented above promises to generate vast amounts of data that need to be processed efficiently and effectively. This opens up a number of data analysis challenges.

A. Data segmentation and validation

We hope to replace the current laborious process of hand-segmenting the image data to identify areas of linguistic interest by an automated, *data-driven* process. A first goal is to enable automatic methods for extracting relevant regions of interest from the image sequences. Our ongoing work focuses on methods to automatically segment and track the real-time MRI data using Kalman snakes and optical flow (Kass *et al.*, 1987; Cootes *et al.*, 1994; Gautama and van Hulle, 2002). At this time, this has allowed us to quantitatively track events during articulation; for example, the open-

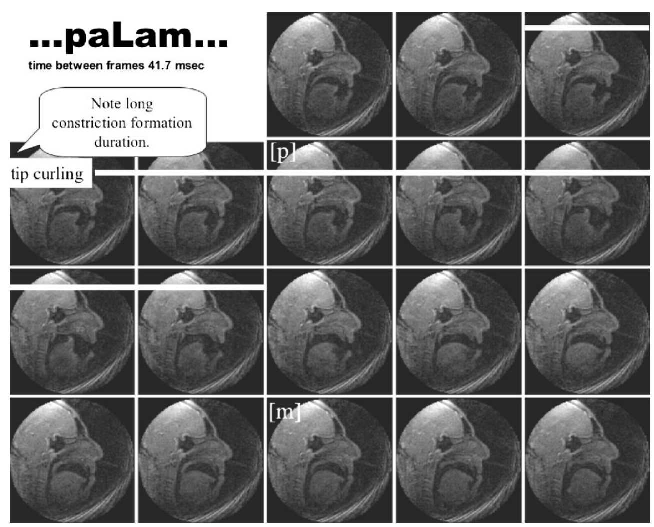

FIG. 8. An example sequence for "...paɭam..." with 41.7 ms between frames. The middle panel in the top row shows the beginning of tongue tip curling toward forming the retroflex lateral; the maximal tongue retroflexion can be seen in the first panel of row 3. Note the formation of the word-initial and word-final labials are also marked in the middle frames of rows 2 and 4, respectively.

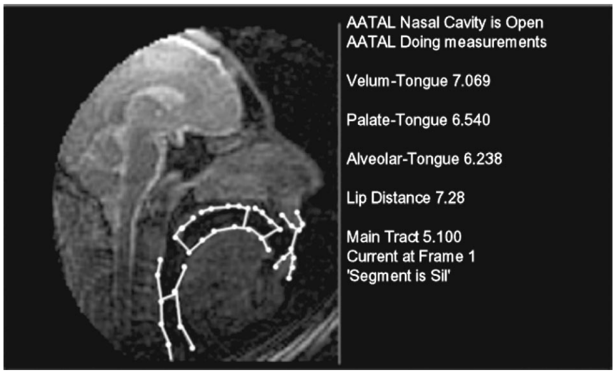

FIG. 9. A snapshot of the automatic tracking method of MR images developed in the study. In the left panel, each group of points connected by line represents a Kalman snake (i.e., contour), which is individually updated through an optical flow method applied to the sequence of images. In this demonstration the midsagittal dimensions (represented by lines connecting the upper and lower walls of the vocal tract) at some positions along the vocal tract are computed and corresponding values are displayed in the right panel. (A) The real-time MRI movie can be downloaded at sail.usc.edu/production/rtmri/jasa2004.

ing and closing of the velic port was detected based on a edge detection and linking method. We present a preliminary movie (frame captured in Fig. 9) that shows automatic airway tracking and measurements (the movie can be downloaded at sail.usc.edu/production/rtmri/jasa2004). Future work will focus on issues related to calibration and validation of such automatic processing schemes. Once validated, these data can be valuable for modeling, such as in performing degree-of-freedom analysis and motion parametization of articulators using PCA applied to the point-tracking data generated by image analysis.

B. Combining data types

Parallel point-tracking (EMMA magnetometer; see Perkell *et al.*, 1992) data collected previously for the same Tamil stimuli and speaker (Narayanan *et al.*, 1999) provide a future opportunity to investigate data alignment strategies for MRI and point-tracking technologies. For example, key region tracking from MRI can be supplemented and/or validated by using magnetometer data. In the future, imagable beads can help coregistration for validation purposes (Byrd *et al.*, 1999). In addition to improving the overall time resolution of the combined data, point-tracking data also provide us with dynamical information in terms of velocity and acceleration for different points on active articulators.

C. New avenues

The ability to acquire real-time images of the speaking vocal tract can provide vital data for computational modeling of the production process that are not currently otherwise available. Intergestural coordination can be evaluated—for example, an understanding of the production of English [r] can be enhanced by the fact that the pharyngeal, oral, and labial components of the consonant can be simultaneously quantified (Alwan *et al.*, 1997), not only for the quality of the constrictions but also for the coordination among the constrictions (note that point-tracking technology, in contrast, cannot be employed in the pharynx). Rhotic mechanisms can be examined cross-linguistically, including rare or poorly described rhotics such as the uvular.

We also feel that this real-time MRI tool can provide heretofore unavailable details that will inform our understanding of interspeaker variability, since individual differences in the morphology of the vocal tract can be directly associated with differences in speech production behavior.

D. Technical improvements on the horizon

We have identified and are pursuing a number of avenues to further improve both the speed and quality of the image acquisition. First, our team is in the process of designing a new neck coil for better SNR for airway region. Second, we are exploring options for the acquisition of (even compromised) audio signals during imaging (this has been done fairly successfully in prior static vocal-tract imaging studies). Finally, a significant effort is being devoted to new pulse sequence designs. In spiral imaging (such as the one reported here), off-resonance causes a blurring of signal (Noll *et al.*, 1991) predominantly at the air–tissue interface. When more than $\pi/4$ of phase is accrued during a readout this blurring is significant (Noll *et al.*, 1992), and this small pervasive blurring can be observed throughout our image sequences, for example, at the tongue surface. Another fast alternative is echo-planar imaging, where off-resonance results in a shifting of signal along the phase encode direction. The amount of shift in pixels is roughly equal to the amount of phase accrued over a readout divided by 2π. In areas of high off-resonance, echo-planar can result in a warping artifact but will retain image sharpness. An additional benefit of echo-planar is that the FOV can be limited in the readout direction using the analog filter, enabling scanning with a smaller FOV. Both of these effects may enable sharper and higher resolution real-time imaging of the upper airway. While preliminary experiments have shown improved spatial resolution (1.56 mm) and improved sharpness, further investigation is needed to establish the performance of echo-planar in the presence of rapid motion with speech.

V. CONCLUSION

Using a new approach, real-time images of the moving vocal tract with MRI are, for the first time, possible at rates that permit meaningful investigations of speech production dynamics ($\geqslant$20 images per second). Future challenges for this technology include improvements in image quality and in quantitative evaluation of continuously varying vocal-tract shaping.

ACKNOWLEDGMENTS

This work was supported by NIH Grant DC 03172. The authors extend their thanks to Phil Hoole and an anonymous reviewer for their helpful comments.

[1]Cinefluorography does permit this but is not useful due to the need to protect human subjects.

Alwan, A., Narayanan, S., and Haker, K. (**1997**). ''Toward articulatory-

acoustic models for liquid consonants based on MRI and EPG data. II. The rhotics,'' J. Acoust. Soc. Am. **101**, 1078–1089.
Baer, T., Gore, J., Gracco, C., and Nye, R. (**1991**). ''Analysis of vocal tract shape and dimensions using magnetic resonance imaging: Vowels,'' J. Acoust. Soc. Am. **90**, 799–828.
Bangayan, P., Alwan, A., and Narayanan, S. (**1996**). ''From MRI and acoustic data to articulatory synthesis: A case study of the lateral approximants in American English,'' in Proceedings of the International Conference. Spoken Language Processing (Philadelphia, PA), pp. 793–796.
Byrd, D., Browman, C. P., Goldstein, L., and Honorof, D. (**1999**). ''Magnetometer and x-ray microbeam comparison,'' in *Proceedings of the 14th International Congress of Phonetic Sciences*, edited by J. J. Ohala, Y. Hasegawa, M. Ohala, D. Granville, and A. C. Bailey (American Institute of Physics, New York), pp. 627–630.
Cootes, T. F., Hill, A., Taylor, C. J., and Haslam, J. (**1994**). ''The use of active shape models for locating structures in medical images,'' Image Vis. Comput. **12**, No. 6, 355–366.
Dang, J., Honda, K., and Suzuki, H. (**1993**). ''MRI measurements and acoustic investigation of the nasal and paranasal cavities.'' J. Acoust. Soc. Am. **94**, No. 3, 1765.
Delattre, P., and Freeman, D. C. (**1968**). ''A dialect study of American r's by x-ray motion picture,'' Linguistics, An International Review, Vol. 44, pp. 29–68.
Demolin, D., Metens, T., and Soquet, A. (**2000**). ''Real time MRI and articulatory coordinations in vowels speech production,'' Proceedings of the 5th Speech Production Seminar, pp. 86–93.
Espy-Wilson, C. Y., Boyce, S. E., Jackson, M. T. T., Narayanan, S., and Alwan, A. (**2000**). ''Acoustic modeling of the American English /r/,'' J. Acoust. Soc. Am. **108**, No. 1, 343–356.
Fant, G. (**1960**). *Acoustic Theory of Speech Production* (Mouton and Co., The Hague).
Foldvik, A., Kristainsen, U., and Kvaerness, J. (**1993**). ''A time-evolving three-dimensional vocal tract model by means of MRI,'' Proceedings of Eurospeech, pp. 557–560.
Gautama, T., and van Hulle, M. M. (**2002**). ''A phase-based approach to the estimation of the optical flow field using spatial filtering,'' IEEE Trans. Neural Netw. **13**(5), 1127–1136.
Giles, S. B., and Moll, K. L. (**1975**). ''Cineflurographic study of selected allophones of English /1/,'' Phonetica **31**, 206–227.
Greenwood, A. R., Goodyear, C. C., and Martin, P. A. (**1992**). ''Measurement of vocal tract shapes using magnetic resonance imaging,'' IEE Proc. I (Commun., Speech Vision) (UK), Vol. 139, No. 6, pp. 553–560.
Holsinger, A., Wright, R. C., Riederer, S. J., Farzaneh, F., Grimm, R. C., and Maier, J. K. (**1990**). ''Real-time interactive magnetic resonance imaging,'' Magn. Reson. Med. **14**, No. 3, 547–553.
Kass, M., Witkin, A., and Terzopoulos, A. D. (**1987**). ''Snake: Active contour model,'' Int. J. Comput. Vis. **1**, 321–331.
Kerr, A. B., Pauly, J. M., Meyer, C. H., and Nishimura, D. G. (**1995**). ''New strategies in spiral MR fluoroscopy,'' Proceedings of SMR, 3rd Annual Meeting (Nice, FR), p. 99.
Kerr, A. B., Pauly, J. M., Hu, B. S., Li, K. C. P., Hardy, C. J., Meyer, C. H., Macovski, A., and Nishimura, D. G. (**1997**). ''Real-time interactive MRI on a conventional scanner,'' Magn. Reson. Med. **38**, 355–367.
Kiritani, S., Tateno, Y., and Iinuma, T. (**1977**). ''Computer tomography of the vocal tract,'' in *Dynamic Aspects of Speech Production*, edited by M. Sawashima and F. S. Cooper (University of Tokyo press, Tokyo), pp. 203–206.
Mády, K., Sader, R., Zimmermann, A., Hoole, P., Beer, A., Zeilhofer, H., and Hannig, C. (**2001**). ''Use of real-time MRI in assessment of consonant articulation before and after tongue surgery and tongue reconstruction.'' in *Speech Motor Control in Normal and Disorderd Speech*, edited by B. Maassen, W. Hulstijn, R. Kent, H. Peters, and P. van Lieshout, 4th International Speech Motor Conference, Nijmegen, pp. 142–145.
Mády, K., Sader, R., Zimmermann, A., Hoole, P. Beer, A., Zeilhofer, H., and Hannig, C. (**2002**). ''Assessment of consonant articulation in glossectomee speech by dynamic MRI,'' Proceedings of ICSLP, Denver.
Meyer, C. H., Hu, B. S., Nishimura, D. G., and Macovski, A. (**1992**). ''Fast spiral coronary artery imaging,'' Magn. Reson. Med. **28**, 202–213.
Moore, C. A. (**1992**). ''The correspondence of vocal tract images with volumes obtained from magnetic resonance images,'' J. Speech Hear. Res. **35**, 1009–1023.
Narayanan, S., and Alwan, A. (**2000**). ''Noise source models for fricative consonants,'' IEEE Trans. Speech Audio Process. **8**, No. 3, 328–344.
Narayanan, S., Alwan, A., and Haker, K. (**1995**). ''An articulatory study of fricative consonants using magnetic resonance imaging,'' J. Acoust. Soc. Am. **98**, 1325–1347.
Narayanan, S., Alwan, A., and Haker, K. (**1997**). ''Toward articulatory-acoustic models for liquid consonants based on MRI and EPG data. I. The laterals,'' J. Acoust. Soc. Am. **101**, 1064–1077.
Narayanan, S., Byrd, D., and Kaun, A. (**1999**). ''Geometry, kinematics, and acoustics of Tamil liquid consonants,'' J. Acoust. Soc. Am. **106**, 1993–2007.
Nayak, K. S., Pauly, J. M., Yang, P. C., Hu, B. S., Meyer, C. H., and Nishimura, D. G. (**2001**). ''Real-time interactive coronary MRA,'' Magn. Reson. Med. **46**, No. 3, 430–435.
Nishimura, D. G., Irarrazabal, P., and Meyer, C. H. (**1995**). ''A velocity k-space analysis of flow effects in echo-planar and spiral imaging,'' Magn. Reson. Med. **33**, No. 4, 549–556.
Noll, D. C., Meyer, C. H., Pauly, J. M., Nishimura, D. G., and Macovski, A. (**1991**). ''A homogeneity correction method for magnetic resonance imaging with time-varying gradients,'' IEEE Trans. Med. Imaging **10**, No. 4, 629–637.
Noll, D. C., Pauly, J. M., Meyer, C. H., Nishimura, D. G., and Macovski, A. (**1992**). ''Deblurring for non-2D Fourier transform magnetic resonance imaging,'' Magn. Reson. Med. **25**, 319–333.
Ong, D., and Stone, M. (**1998**). ''Three-dimensional vocal tract shapes in /r/ and /l/: A study of MRI, ultrasound, electropalatography, and acoustics,'' Phonoscope **1**, 1–13.
Perkell, J. S. (**1969**), ''Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study,'' (Cambridge, The MIT Press).
Perkell, J. S., Cohen, M., Svirsky, M., Matthies, M., Garabieta, I., and Jackson, M. (**1992**). ''Electromagnetic midsagittal articulometer systems for transducing speech articulatory movements,'' J. Acoust. Soc. Am. **92**, No. 6, 3078–3096.
Santos, J. M., Wright, G., Yang, P. C., and Pauly, J. M. (**2002**). ''Adaptive architecture for real-time imaging systems,'' Proceedings of ISMRM, 10th Annual Meeting (Honolulu, HI), p. 468.
Schenk, J. F. (**1996**). ''The role of magnetic susceptibility in magnetic resonance imaging: MRI magnetic compatibility of the first and second kinds,'' Med. Phys. **23**, No. 6, 815–850.
Shadle, C. H., Mohammad, M., Carter, J., and Jackson, P. J. B. (**1999**). ''Dynamic magnetic resonance imaging: New tools for speech research,'' Proceedings of the 14th Int. Cong. Phon. Sci., pp. 623–626.
Shellock, H. G., Schatz, C. J., Julien, P. M., Silverman, J. M., Steinberg, F., Foo, T. K., Hopp, K. L., and Westbrook, P. R. (**1992**). ''Dynamic study of the upper airway with ultrafast spoiled GRASS MR imaging,'' J. Magn. Reson Imaging **2**, 103–107.
Stone, M. (**1990**). ''A three-dimensional model of tongue movement based on ultrasound and x-ray microbeam data,'' J. Acoust. Soc. Am. **87**, 2207–2217.
Stone, M., Faber, A., Raphael, L. J., and Shawker, T. H. (**1992**). ''Cross-sectional tongue shapes and linguopalatal contact patterns in [s], [ʃ], and [l],'' J. Phonetics **20**, No. 2, 253–270.
Story, B. (**1995**). ''Physiologically based speech simulation using an enhanced wave-reflection model of the vocal tract,'' Ph.D. dissertation, University of Iowa.
Story, B., Titze, I., and Hoffman, E. (**1996**). ''Vocal tract area functions from magnetic resonance imaging,'' J. Acoust. Soc. Am. **100**, 537–554.
Subtelny, J. D., Oya, N., and Subtelny, J. D. (**1972**). ''Cineradiographic study of sibilants,'' Folia Phoniatr. **24**, 30–50.
Sulter, A., Miller, D., Wolf, R., Schutte, H., Wit, H., and Mooyaart, E. (**1992**). ''On the relation between the dimensions and resonance characteristics of the vocal tract: A study with MRI,'' Magn. Reson. Imaging **10**, 365–373.
Watkin, K. L., and Rubin, J. M. (**1989**). ''Pseudo-three-dimensional reconstruction of ultrasonic images of the tongue,'' J. Acoust. Soc. Am. **85**, 496–499.