

Viewing speech in action: speech articulation videos in the public domain that demonstrate the sounds of the International Phonetic Alphabet (IPA)

Satsuki Nakai, David Beavan, Eleanor Lawson, Grégory Leplâtre, James M. Scobbie & Jane Stuart-Smith

To cite this article: Satsuki Nakai, David Beavan, Eleanor Lawson, Grégory Leplâtre, James M. Scobbie & Jane Stuart-Smith (2018) Viewing speech in action: speech articulation videos in the public domain that demonstrate the sounds of the International Phonetic Alphabet (IPA), *Innovation in Language Learning and Teaching*, 12:3, 212-220, DOI: [10.1080/17501229.2016.1165230](https://doi.org/10.1080/17501229.2016.1165230)

To link to this article: <https://doi.org/10.1080/17501229.2016.1165230>



© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 11 Apr 2016.



[Submit your article to this journal](#)



Article views: 12791



[View related articles](#)









[View Crossmark data](#)



Citing articles: 2 [View citing articles](#)

Viewing speech in action: speech articulation videos in the public domain that demonstrate the sounds of the International Phonetic Alphabet (IPA)

Satsuki Nakai^a , David Beavan^b , Eleanor Lawson^a , Grégory Leplâtre^c ,
James M. Scobbie^a  and Jane Stuart-Smith^d 

^aClinical Audiology, Speech and Language (CASL) Research Centre, Queen Margaret University, Musselburgh, UK; ^bCentre for Digital Humanities (UCLDH), Faculty of Arts and Humanities, University College London, London, UK; ^cSchool of Computing, Edinburgh Napier University, Edinburgh, UK; ^dEnglish Language, University of Glasgow, Glasgow, UK

ABSTRACT

In this article, we introduce recently released, publicly available resources, which allow users to watch videos of hidden articulators (e.g. the tongue) during the production of various types of sounds found in the world's languages. The articulation videos on these resources are linked to a clickable International Phonetic Alphabet chart ([International Phonetic Association. 1999. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press]), so that the user can study the articulations of different types of speech sounds systematically. We discuss the utility of these resources for teaching the pronunciation of contrastive sounds in a foreign language that are absent in the learner's native language.

ARTICLE HISTORY

Received 14 December 2015
Accepted 25 February 2016

KEYWORDS

Pronunciation; foreign language; articulation video; IPA; MRI; ultrasound

1. Introduction

It is extremely difficult, if not impossible, to fully master the phonology of a non-native language, unless exposure to that language starts in early childhood (Seliger 1978). Yet, pronunciation is an aspect that many language learners aspire to master (Fraser 2010). Ironically, it is also an area that received little attention in the language classroom in the recent past, as well as an area that many teachers lack confidence in teaching (Foote 2015).

Pronunciation was a major concern of approaches such as the Reform Movement at the turn of the twentieth century, and the Audiolingual Method in the 1940–1950s. The later marginalisation of pronunciation instruction can be ascribed to several factors. Notably, the Communicative Language Teaching, which emphasised meaning over form, gained popularity in the 1970s. At around the same time, the efficacy of classroom pronunciation instruction was also questioned (e.g. Suter 1976 as cited in Ketabi and Saeb 2015). During the past decade, however, calls for pronunciation instruction have increased. This was preceded by a gradual spread of a view that 'intelligible pronunciation is an essential part of communicative competence' (Morley 1991, 488), and the emergence of evidence that pronunciation instruction is beneficial (Derwing and Munro 2005).

Along with the renewed interest in pronunciation instruction came new perspectives. Intelligibility and comprehensibility replaced native-like pronunciation as the official goal of the language classroom. The new goal is motivated by feasibility, practicality, and ideology, the latter two particularly

CONTACT Satsuki Nakai  satsuki@ovod.net

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

for English, today's global lingua franca (Jenkins 2002; Jindapitak 2015; Thomson and Derwing 2015). In attaining intelligibility and comprehensibility, both individual sounds and suprasegmental features of speech (e.g. stress and intonation) are now considered important (Lee, Jang, and Plonsky 2014; Foote 2015; Ketabi and Saeb 2015). To what extent pronunciation instruction, which focuses on form, can be integrated into the speaking and listening components within a communicative framework remains to be seen (Isaacs 2009; Foote 2015).

Whatever the perspective, few would disagree that an inability to adequately produce contrastive sounds, or phonemes, in the target language (e.g. *rice* vs. *lice*) can hinder communication (Jenkins 2002). While articulatory instructions are effective in teaching the pronunciation of phonemes absent in the learner's native language (e.g. Catford and Pisoni 1970), traditional articulation-based descriptions of speech sounds may be daunting for language teachers as well as learners who are not accustomed to articulatory phonetics (Yule 1990).

In this article, we introduce recently released, web-based resources that would help such users grasp how different types of speech sounds are produced through videos of hidden articulators (e.g. the tongue) in action. These videos are linked to clickable International Phonetic Alphabet (IPA) charts (International Phonetic Association 1999), so that the users can study the articulation of various speech sounds systematically. The resources are accessible to anyone with Internet access, for free.

The rest of the article is organised as follows. Section 2 gives a technical overview of some data acquisition techniques used to capture hidden speech articulators (Section 2.1) and the IPA chart (Section 2.2). Section 3 introduces resources providing articulation videos linked to the IPA chart. Section 4 provides notes regarding the usage of these resources for improving pronunciation in a foreign language. The web-based resources can be accessed by clicking the URLs (Uniform Resource Locators) after the name of each resource, where it is discussed in the article text. The resources' URLs are also listed in Section 5. Section 6 provides a glossary of some phonological/phonetic terms used in this article; in the article text these terms are given in small capital letters on their first use.

2. Technical overview

2.1. Articulatory data acquisition techniques

When we speak, multiple articulators (e.g. **VOCAL FOLDS**, lips, **VELUM**, tongue, jaw) are delicately coordinated to produce strings of sounds that make up words, phrases, and sentences. Many of these articulators are difficult to observe during speech production. The X-ray technique was used in the past to study speech production, as X-rays make all articulators visible (Stone 2013). Though X-rays shed tremendous light on speech production, their use is now largely restricted to medical imaging because of increased awareness of associated health risks.

In place of X-rays, magnetic resonance imaging (MRI) is now widely used to study the **VOCAL TRACT** shape during speech production. MRI utilises radio signals produced by the hydrogen atoms in water and fat in our body, in response to a strong magnetic field created nearby. Figure 1(a) shows the mid-sagittal plane (profile view) of a female speaker's head, captured by MRI during her production of an **ALVEOLAR PLOSIVE** consonant [d], as in *dine*. As the figure shows, MRI characterises soft tissues well, and the overall vocal tract shape can be clearly seen, although MRI cannot easily image teeth, which contain little water or fat. The latter aspect is a drawback of MRI applications in speech research, as teeth play an important role when producing sounds like [f] and [θ], as in *fin* and *thin*. Furthermore, as the weak radio signals elicited during MRI scans need to be summed over time, it is difficult to obtain high-quality images at high frame rates, required to capture the rapid movements of some speech articulators, notably, the tongue. Other drawbacks include loud scanning noise, which interferes with audio recording (NessAiver et al. 2006).

Tongue movements during speech production are instead often studied using ultrasound tongue imaging (UTI). UTI capitalises on the fact that sound waves pass well through soft tissue but are

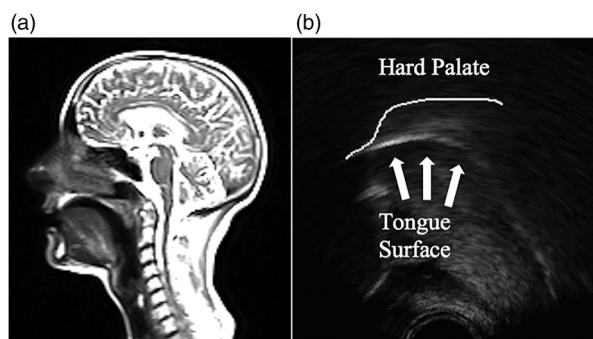


Figure 1. Midsagittal (a) MRI and (b) UTI scans of a production of [d] during closure. The white curve descending towards the lower right corner in Figure 1(b) corresponds to the tongue surface; another white line indicating the **HARD PALATE** was added later to the UTI scan.

reflected at tissue-air boundaries; real-time images of the upper surface of the tongue can thus be viewed during speech production by placing beneath the speaker's chin, a probe that emits and receives ultrahigh frequency sound waves. The strengths and weaknesses of UTI are almost complementary to those of MRI. As Figure 1(b) shows, UTI mainly shows the tongue surface only, and how the tongue surface relates to the rest of the vocal tract is unclear. Additionally, the tongue tip is often invisible in UTI data because of an air pocket underneath, or a shadow cast by the jaw bone. However, with UTI it is easier to achieve frame rates high enough to capture most tongue motions during speech production, with many modern ultrasound machines capable of scanning at around 90 fps (frames per second) (Stone 2013). Other advantages to UTI when compared with MRI include relative ease with which high-quality audio recordings can be obtained during data acquisition, as well as relatively low cost and convenience.

2.2. The IPA chart

The IPA consists of a universal set of symbols representing distinctive sounds of the world's languages, and is used to show pronunciation in many dictionaries (International Phonetic Association 1999). The IPA chart consists of several sections such as vowels, **PULMONIC** consonants, and non-pulmonic consonants.

The IPA chart can be a useful tool for teaching the basics of speech production, as it shows at a glance commonalities and differences between the articulations of various speech sounds. The vowels section of the chart, shown in Figure 2, arranges vowels along three dimensions: tongue height, tongue backness, and lip rounding. Figure 3 gives a partial view of the pulmonic consonants section. Each column of the table represents a place of articulation from **BILABIAL** to **GLOTTAL** (from the leftmost to the rightmost column), and each row a manner of articulation (e.g. plosive, **NASAL**, **TRILL**). It is easy to see in the table, for example, that [d] and [n] (as in *dine* and *nine*) are produced at similar places of articulation but with different manners of articulation (see International Phonetic Association 1999 for details).

Figure 4(a)–(b) shows the midsagittal plane of a female speaker's head, captured by MRI during her production of an alveolar plosive consonant [d] vs. an alveolar nasal consonant [n]. The two consonants are produced with similar tongue shapes, reflecting the same place of articulation. However, the velum is raised for [d] to build up the air pressure behind the tongue, and lowered for [n] to let the air stream pass through the nose. It would be helpful if one could access videos of hidden articulators in action, when learning the articulations of phonemes absent in one's native language. Such organised resources are now publicly available.

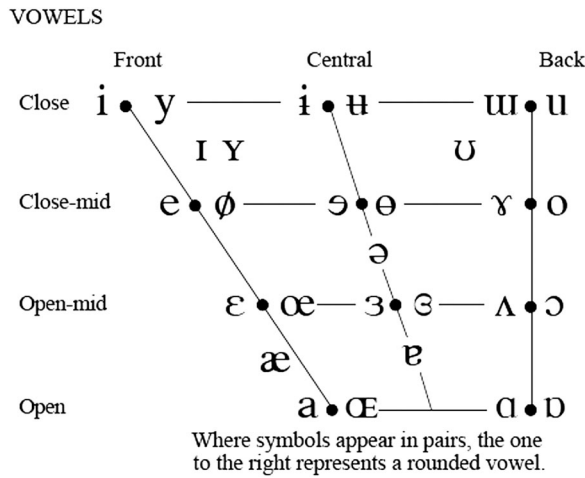


Figure 2. Vowels section of the IPA chart. IPA chart, <http://www.internationalphoneticassociation.org/content/ipa-chart>, available under a Creative Commons Attribution-Sharealike 3.0 Unported License. Copyright © 2005 International Phonetic Association.

CONSONANTS (PULMONIC)

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	R
Plosive	p b			t d		
Nasal	m	ɱ		n		
Trill	ʙ			r		
Tap or Flap		ⱱ		ɾ		

Figure 3. Partial view of the pulmonic consonants section of the IPA chart.

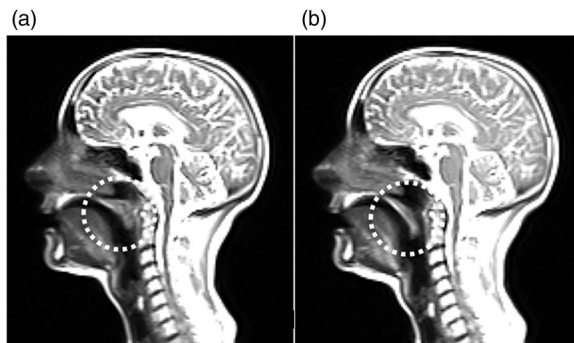


Figure 4. Midsagittal MRI scans of a production of (a) an alveolar plosive [d] vs. (b) an alveolar nasal [n], during closure. The raised vs. lowered velum can be seen in the areas marked with dotted circles. Figure 4(a) is a reproduction of Figure 1(a).

3. Articulation video resources

The 'IPA Charts' section of the *Seeing Speech* website (Lawson et al. 2015b) <http://www.seeing-speech.ac.uk>, hosted by the University of Glasgow, provides a clickable IPA chart, which allows the user to access the articulation videos of desired sounds by clicking sound symbols on the chart. The

articulation videos on *Seeing Speech* come in three forms: MRI videos, UTI videos, and animated heads. The MRI videos feature a female phonetician and are synchronised with temporally matched, high-quality audio recordings of the same phonetician made in a quiet recording condition. The UTI videos, which feature the same female phonetician and a male phonetician, are played at original and then half speed.

The animated heads were created by combining the image of the entire vocal tract from the female phonetician's MRI data, with her tongue data from UTI and lip data from a video camera, both captured at better temporal and spatial resolutions than MRI. The resulting animations were sampled at 24 fps, at a higher frame rate than the 6.5 fps of the MRI data. Figure 5(a)–(c) outlines the animation creation process. As Figure 5(c) shows, the animated head has the upper and lower teeth; the location of the biting edge of the teeth was estimated from an MRI recording of an **INTER-DENTAL** articulation (where the incisors indent the tongue's surface). Individual videos on this website are also available on YouTube's ArticulatoryIPA channel <https://www.youtube.com/playlist?list=UUuOKJqD00W2EiC3DHmOuu0g>.

Another useful resource with a clickable IPA chart is *rtMRI IPA Chart* http://sail.usc.edu/span/rtmri_ipa/index.html, created by the Speech Production and Articulation kKnowledge Group, University of Southern California. This website was recently updated to provide high-frame-rate MRI videos of two female and two male phoneticians, synchronised with noise-cancelled audio recordings made simultaneously with the MRI recordings. The videos are sampled at 83 fps, using a sliding window technique to reconstruct MRI scans at a much higher frame rate than the rate at which the data were acquired (Narayanan et al. 2004). Along with the sounds on the IPA chart, the website provides MRI videos showing the production of some English words, sentences, and passages by phoneticians from the US.

The articulation videos on *Seeing Speech* and *rtMRI IPA Chart* can be played back without a plugin (e.g. QuickTime), that is, they can be viewed on smartphones and tablets as well as computers. Accessibility on devices used daily by the learners is potentially important, as improvement in pronunciation typically does not occur overnight (Thomson and Derwing 2015).

Finally, though the platform is limited to iOS devices (e.g. iPad, iPhone), *iPA Phonetics* (Coey, Esling, and Moisik 2014) allows the user to view videos of the **LARYNX** linked to an expanded IPA chart, and study **LARYNGEAL** articulations for various types of speech sounds found in the world's languages (Esling, Moisik, and Coey 2015). The application is free on the Apple Store.

4. Notes on usage of the resources

These recently released, publicly available resources provide articulation videos, which we hope will be useful when teaching/learning the pronunciation of phonemes absent in the learner's native language. We consider below how these resources should be used.

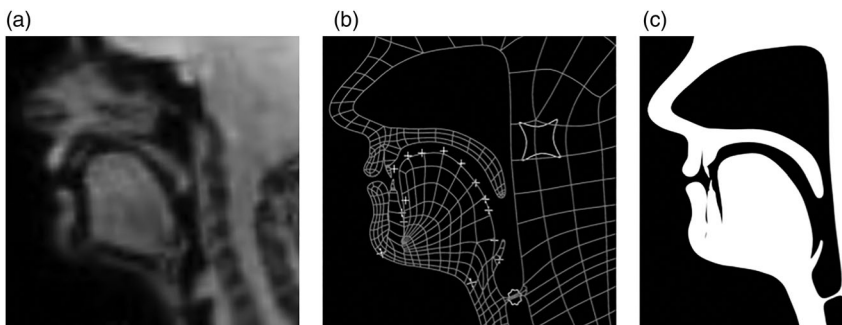


Figure 5. (a) Midsagittal MRI scan of the model-talker's head, (b) animation rig created from the MRI scan (white symbols show control points), and (c) drawing used for animation.

First, as explained earlier, the IPA groups contrastive sounds of the world's languages into a set of universal phonetic categories. Thus, the articulation videos and associated audio recordings in the IPA-based resources broadly reflect different ways in which sounds represented by different symbols are produced *independent of* language; they necessarily do not tell us about more subtle differences in the realisation of any given sound category that exist *between* languages or language varieties (Ladefoged 1987). Consequently, these resources can be used to teach basic articulatory principles underlying the production of phonemes absent in the learner's native language, but not to teach cross-linguistic or cross-dialectal differences in the way a sound category is produced.¹

Second, their utility as audiovisual aids in formal articulatory instructions aside, how best to use these resources for drill-type pronunciation training is yet to be explored. There are some positive signs from studies reporting improved pronunciation of non-native phonemes through training with animated heads revealing the workings of internal speech articulators, and the learners' appreciation of such input (Massaro and Light 2003; Kröger et al. 2010; Dey 2012; Wang, Hueber, and Badin 2014). Nevertheless, it is surprisingly difficult to find studies unambiguously demonstrating *additional* benefits from visual presentation of hidden speech articulators, when compared with images of a model speaker's face.²

While this may be partially ascribed to the paucity of studies directly addressing the question, it is worth bearing in mind that many late learners bring to language learning well-developed skills to extract articulatory information from the speaker's face (Hazan et al. 2005), but not from visual presentations of internal articulators. The following summarises the findings of studies on our ability in native language to extract articulatory information from visual displays of complete vs. cutaway heads presented with degraded acoustic information (Grauwinkel, Dewitt, and Fagel 2007; Wik and Engwall 2008; Badin et al. 2010). First, we are generally better at extracting articulatory information from videos of complete heads than cutaway heads. With training, many of us can learn to use articulatory information conveyed by cutaway heads, but individual differences in this ability are large. Second, learning to interpret a cutaway head is easier when the visual information is presented alone, than when the visual and acoustic information are presented bimodally.

Thus, familiarisation with videos of hidden articulators is perhaps required before learners can fully benefit from the videos, with some learners requiring more time and training (hence the potential importance of accessibility of these resources on smartphones and tablets). Furthermore, familiarisation and/or training may be more effective if carried out in stages, for example, by first presenting the videos without acoustic information and later with acoustic information (as in, for example, Bernhardt et al. 2008).³ These recommendations are only tentative, as they are based on an assumption that native language speech perception in noise can be compared to the situation faced by late language learners, who cannot hear all the sound contrasts in the target language. We welcome future research on the effective use of videos of hidden speech articulators in pronunciation training.

5. URLs of articulatory resources

Seeing Speech: <http://www.seeingspeech.ac.uk>

ArticulatoryIPA channel on YouTube: <https://www.youtube.com/playlist?list=UUuOKJqD00W2EiC3DHmOuu0g>

rtMRI IPA Chart: http://sail.usc.edu/span/rtmri_ipa/index.html

Dynamic Dialects: <http://www.dynamicdialects.ac.uk>

6. Glossary of phonological/phonetic terms

ALVEOLAR (A sound) associated with the placement of the tongue against/near the **HARD PALATE**, just behind the teeth.

BILABIAL (A sound) produced using the upper and lower lips.

GLOTTAL (A sound) produced with the glottis, that is, the opening between the **VOCAL FOLDS**.

HARD PALATE The bony, front part of the roof of the mouth.

INTERDENTAL (A sound) associated with the placement of the tongue tip between the upper and lower front teeth.

LARYNGEAL Relating to LARYNX.

LARYNX A complex organ, which houses the VOCAL FOLDS. It sits in the neck, below the tongue root and above the windpipe.

NASAL (A sound) produced with passage of air through the nose.

PLOSIVE (A sound) produced by complete closure of oral passage followed by a sudden release of higher pressure air.

PULMONIC (A sound) produced using the air from the lungs.

TRILL (A sound) produced by a rapid vibration of an articulator against another.

VELUM The soft, back part of the roof of the mouth.

VOCAL FOLDS Two folds of tissue in the LARYNX. They vibrate to produce voiced sounds.

VOCAL TRACT The branching passage between the LARYNX and the mouth and nose.

Notes

1. Readers interested in articulation videos of words and phrases produced with various regional accents of English are referred to the *Dynamic Dialects* website (Lawson et al. 2015a) <http://www.dynamicdialects.ac.uk>.
2. This should not be confused with reports of positive training outcomes using visual biofeedback provided by, for example, electropalatography or UTI for pronunciation training in a foreign language (Schmidt 2012; Wu et al. 2015) and speech therapy (Gibbon et al. 1996; Bernhardt et al. 2008).
3. Relatedly, the benefits of visual presentation of internal articulators appear dramatic in speech production training of people with hearing loss (Massaro 2003).

Acknowledgements

We thank Asterios Toutios of the Speech Production and Articulation kKnowledge Group, University of Southern California, for his input regarding *rtMRI IPA*.

Funding

This work was supported by the Arts and Humanities Research Council (AHRC) [grant number AH/L010380/1]. The MRI scans in this article were obtained with the financial support of NHS Research Scotland (NRS), using a 3T Siemens Verio MRI system at Edinburgh Clinical Research Facility, under the supervision of Gillian Macnaught and Scott Semple.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Satsuki Nakai is a phonetician affiliated to the Clinical Audiology, Speech and Language (CASL) Research Centre, Queen Margaret University, UK.

David Beavan is Associate Director for Research at University College London Centre for Digital Humanities (UCLDH), UK.

Eleanor Lawson is a researcher in articulatory phonetics at CASL, Queen Margaret University, UK.

Grégory Leplâtre is a lecturer at the School of Computing, Edinburgh Napier University, UK.

James M. Scobbie is Professor of Speech Sciences at CASL, Queen Margaret University, UK.

Jane Stuart-Smith is Professor of Phonetics and Sociolinguistics in English Language, University of Glasgow, UK.

ORCID

Satsuki Nakai  <http://orcid.org/0000-0002-8540-2652>
 David Beavan  <http://orcid.org/0000-0002-0347-6659>
 Eleanor Lawson  <http://orcid.org/0000-0003-1812-387X>
 Grégory Leplâtre  <http://orcid.org/0000-0002-8857-6367>
 James M. Scobbie  <http://orcid.org/0000-0003-4509-6782>
 Jane Stuart-Smith  <http://orcid.org/0000-0001-7400-9436>

References

- Badin, Pierre, Yuliya Tarabalka, Frédéric Elisei, and Gérard Bailly. 2010. "Can You 'Read' Tongue Movements? Evaluation of the Contribution of Tongue Display to Speech Understanding." *Speech Communication* 52 (6): 493–503.
- Bernhardt, B. May, Penelope Bacsfalvi, Marcy Adler-Bock, Reiko Shimizu, Audrey Cheney, Nathan Giesbrecht, Maureen O'connell, Jason Sirianni, and Bosko Radanov. 2008. "Ultrasound as Visual Feedback in Speech Habilitation: Exploring Consultative Use in Rural British Columbia, Canada." *Clinical Linguistics and Phonetics* 22 (2): 149–162.
- Catford, John C., and David B. Pisoni. 1970. "Auditory vs. Articulatory Training in Exotic Sounds." *Modern Language Journal* 54 (7): 477–481.
- Coe, Christopher, John H. Esling, and Scott R. Moisiuk. 2014. *iPA Phonetics* (Version 1). Victoria: Department of Linguistics, University of Victoria.
- Derwing, Tracey M., and Murray J. Munro. 2005. "Second Language Accent and Pronunciation Teaching: A Research-Based Approach." *TESOL Quarterly* 39 (3): 379–397.
- Dey, Priya. 2012. "Visual Speech in Technology-Enhanced Learning." PhD diss., University of Sheffield, Sheffield.
- Esling, John H., Scott R. Moisiuk, and Christopher Coe. 2015. "iPA Phonetics: Multimodal iOS Application for Phonetics Instruction and Practice." Proceedings of the 18th international congress of Phonetic Sciences. Glasgow, UK.
- Foot, Jennifer Ann. 2015. *Pronunciation Pedagogy and Speech Perception: Three Studies*. Montreal: Concordia University.
- Fraser, Helen. 2010. "Cognitive Theory as a Tool for Teaching Second Language Pronunciation." In *Fostering Language Teaching Efficiency through Cognitive Linguistics*, edited by Sabine de Knop, Frank Boers, and Teun de Rycker, 357–379. Berlin: Mouton de Gruyter.
- Gibbon, Fiona, William Hardcastle, Hilary Dent, and Fiona Nixon. 1996. "Types of Deviant Sibilant Production in a Group of School-Aged Children, and Their Response to Treatment Using Electropalatography (EPG)." In *Studies in Speech Pathology and Clinical Linguistics: Advances in Clinical Phonetics*, edited by Martin J. Ball and Martin Duckworth, 115–149. Amsterdam: John Benjamins.
- Grauwinkel, Katja, Britta Dewitt, and Sascha Fagel. 2007. "Visual Information and Redundancy Conveyed by Internal Articulator Dynamics in Synthetic Audiovisual Speech." In *Proceedings of Interspeech 2007*, 706–709. Antwerp, Belgium.
- Hazan, Valerie, Anke Sennema, Midori Iba, and Andrew Faulkner. 2005. "Effect of Audiovisual Perceptual Training on the Perception and Production of Consonants by Japanese Learners of English." *Speech Communication* 47 (3): 360–378.
- International Phonetic Association. 1999. *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.
- Isaacs, Talia. 2009. "Integrating Form and Meaning in L2 Pronunciation Instruction." *TESL Canada Journal* 27 (1): 1–12.
- Jenkins, Jennifer. 2002. "A Sociolinguistically Based, Empirically Researched Pronunciation Syllabus for English as an International Language." *Applied Linguistics* 23 (1): 83–103.
- Jindapitak, Naratip. 2015. "English as a Lingua Franca: Learners' Views on Pronunciation." *Electronic Journal of Foreign Language Teaching* 12 (2): 260–275.
- Ketabi, Saeed, and Fateme Saeb. 2015. "Pronunciation Teaching: Past and Present." *International Journal of Applied Linguistics and English Literature* 4 (5): 182–189.
- Kröger, Bernd J., Peter Birkholz, Rüdiger Hoffmann, and Helen Meng. 2010. "Audiovisual Tools for Phonetic and Articulatory Visualization in Computer-Aided Pronunciation Training." In *Development of Multimodal Interfaces: Active Listening and Synchrony*, edited by Anna Esposito, Nick Campbell, Carl Vogel, Amir Hussain, and Anton Nijholt, 337–345. Berlin: Springer.
- Ladefoged, Peter. 1987. "Updating the Theory." *Journal of the International Phonetic Association* 17 (1): 10–14.
- Lawson, Eleanor, Jane Stuart-Smith, James M. Scobbie, and Satsuki Nakai. 2015a. "Dynamic Dialects: An Articulatory Web Resource for the Study of Accents." <http://www.dynamicdialects.ac.uk/>.
- Lawson, Eleanor, Jane Stuart-Smith, James M. Scobbie, and Satsuki Nakai. 2015b. "Seeing Speech: An Articulatory Web Resource for the Study of Phonetics." <http://www.seeingsspeech.ac.uk/>.
- Lee, Junkyu, Juhyun Jang, and Luke Plonsky. 2014. "The Effectiveness of Second Language Pronunciation Instruction: A Meta-analysis." *Applied Linguistics* 36 (3): 345–366.

- Massaro, Dominic W. 2003. "A Computer-Animated Tutor for Spoken and Written Language Learning." In *Proceedings of the 5th International Conference on Multimodal Interfaces*, 172–175. Vancouver: ACM Press.
- Massaro, Dominic W., and Joanna Light. 2003. "Read My Tongue Movements: Bimodal Learning to Perceive and Produce Non-Native Speech /r/ and /l/." *Proceedings of Eurospeech (Interspeech)*, Geneva, Switzerland.
- Morley, Joan. 1991. "The Pronunciation Component in Teaching English to Speakers of Other Languages." *TESOL Quarterly* 25 (3): 481–520.
- Narayanan, Shrikanth, Krishna Nayak, Sungbok Lee, Abhinav Sethy, and Dani Byrd. 2004. "An Approach to Real-Time Magnetic Resonance Imaging for Speech Production." *The Journal of the Acoustical Society of America* 115 (4): 1771–1776.
- NessAiver, Moriel S., Maureen Stone, Vijay Parthasarathy, Yuvi Kahana, and Alex Paritsky. 2006. "Recording High Quality Speech During Tagged Cine-MRI Studies Using a Fiber Optic Microphone." *Journal of Magnetic Resonance Imaging* 23 (1): 92–97.
- Schmidt, Anna Marie. 2012. "Effects of EPG Treatment for English Consonant Contrasts on L2 Perception and Production." *Clinical Linguistics and Phonetics* 26 (11–12): 909–925.
- Seliger, Herbert W. 1978. "Implications of a Multiple Critical Periods Hypothesis for Second Language Learning." In *Second Language Acquisition Research: Issues and Implications*, edited by William C. Ritchie, 11–19. New York: Academic Press.
- Stone, Maureen. 2013. "Laboratory Techniques for Investigating Speech Articulation." In *The Handbook of Phonetic Sciences*, edited by William J. Hardcastle, John Laver, and Fiona Gibbon, 2nd ed., 9–38. Oxford: Wiley-Blackwell.
- Thomson, Ron I., and Tracey M. Derwing. 2015. "The Effectiveness of L2 Pronunciation Instruction: A Narrative Review." *Applied Linguistics* 36: 326–344.
- Wang, Xiaou, Thomas Hueber, and Pierre Badin. 2014. "On the Use of an Articulatory Talking Head for Second Language Pronunciation Training: The Case of Chinese Learners of French." In *Proceedings of 10th International Seminar on Speech Production*, 449–452. Cologne.
- Wik, Preben, and Olov Engwall. 2008. "Can Visualization of Internal Articulators Support Speech Perception?" In *Proceedings of Interspeech*, 2627–2630. Brisbane.
- Wu, Yaru, Cédric Gendrot, Pierre Hallé, and Martine Adda-Decker. 2015. "On Improving the Pronunciation of French /r/ in Chinese Learners by Using Real-Time Ultrasound Visualization." In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow.
- Yule, G. 1990. "Review of Teaching English Pronunciation; Current Perspectives on Pronunciation: Practices Anchored in Theory; and Teaching Pronunciation: English Rhythm and Stress." *System* 18: 107–111.