

Publications by Korin Richmond

- [1] R. Dall, S. Brognaux, K. Richmond, C. Valentini-Botinhao, G. E. Henter, J. Hirschberg, and J. Yamagishi, “Testing the consistency assumption: pronunciation variant forced alignment in read and spontaneous speech synthesis,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Mar. 2016, pp. 5155–5159. [\[pdf\]](#)
- [2] Q. Hu, J. Yamagishi, K. Richmond, K. Subramanian, and Y. Stylianou, “Initial investigation of speech synthesis based on complex-valued neural networks,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Mar. 2016, pp. 5630–5634. [\[pdf\]](#)
- [3] K. Richmond and S. King, “Smooth talking: articulatory join costs for unit selection,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Mar. 2016, pp. 5150–5154. [\[pdf\]](#)
- [4] A. Hewer, I. Steiner, T. Bolkart, S. Wuhner, and K. Richmond, “A statistical shape space model of the palate surface trained on 3D MRI scans of the vocal tract,” in *Proceedings of the 18th International Congress of Phonetic Sciences*, The Scottish Consortium for ICPhS 2015, Ed., Glasgow, United Kingdom, Aug. 2015, ISBN: 978-0-85261-941-4. [\[pdf\]](#)
- [5] A. Hewer, S. Wuhner, I. Steiner, and K. Richmond, “Tongue mesh extraction from 3D MRI data of the human vocal tract,” in *Perspectives in Shape Analysis*, M. Breu, A. M. Bruckstein, P. Maragos, and S. Wuhner, Eds., ser. Mathematics and Visualization. Springer, 2015, (in press).
- [6] Q. Hu, Y. Stylianou, R. Maia, K. Richmond, and J. Yamagishi, “Methods for applying dynamic sinusoidal models to statistical parametric speech synthesis,” in *Proc. ICASSP*, Brisbane, Australia, Apr. 2015. [\[pdf\]](#)
- [7] Q. Hu, Z. Wu, K. Richmond, J. Yamagishi, Y. Stylianou, and R. Maia, “Fusion of multiple parameterisations for DNN-based sinusoidal speech synthesis with multi-task learning,” in *Proc. Interspeech*, Dresden, Germany, Sep. 2015. [\[pdf\]](#)
- [8] K. Richmond, Z. Ling, and J. Yamagishi, “The use of articulatory movement data in speech synthesis applications: an overview - application of articulatory movements using machine learning algorithms [invited review],” *Acoustical Science and Technology*, vol. 36, no. 6, pp. 467–477, 2015. DOI: [10.1250/ast.36.467](#).
- [9] K. Richmond, J. Yamagishi, and Z.-H. Ling, “Applications of articulatory movements based on machine learning,” *Journal of the Acoustical Society of Japan*, vol. 70, no. 10, pp. 539–545, 2015.
- [10] J. Cabral, K. Richmond, J. Yamagishi, and S. Renals, “Glottal spectral separation for speech synthesis,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 2, pp. 195–208, Apr. 2014. DOI: [10.1109/JSTSP.2014.2307274](#).
- [11] Q. Hu, Y. Stylianou, R. Maia, K. Richmond, J. Yamagishi, and J. Latorre, “An investigation of the application of dynamic sinusoidal models to statistical parametric speech synthesis,” in *Proc. Interspeech*, Singapore, Sep. 2014, pp. 780–784. [\[pdf\]](#)

- [12] Q. Hu, Y. Stylianou, K. Richmond, R. Maia, J. Yamagishi, and J. Latorre, “A fixed dimension and perceptually based dynamic sinusoidal model of speech,” in *Proc. ICASSP*, Florence, Italy, May 2014, pp. 6311–6315. [\[pdf\]](#)
- [13] M. Astrinaki, A. Moinet, J. Yamagishi, K. Richmond, Z.-H. Ling, S. King, and T. Dutoit, “Mage - reactive articulatory feature control of HMM-based parametric speech synthesis,” in *8th ISCA Workshop on Speech Synthesis*, Barcelona, Spain, Aug. 2013, pp. 227–231. [\[pdf\]](#)
- [14] —, “Mage-HMM-based speech synthesis reactively controlled by the articulators,” in *8th ISCA Workshop on Speech Synthesis*, Barcelona, Spain, Aug. 2013, p. 243. [\[pdf\]](#)
- [15] C. Geng, A. Turk, J. M. Scobbie, C. Macmartin, P. Hoole, K. Richmond, A. Wrench, M. Pouplier, E. G. Bard, Z. Campbell, C. Dickie, E. Dubourg, W. Hardcastle, E. Kainada, S. King, R. Lickley, S. Nakai, S. Renals, K. White, and R. Wiegand, “Recording speech articulation in dialogue: evaluating a synchronized double electromagnetic articulography setup,” *Journal of Phonetics*, vol. 41, no. 6, pp. 421–431, 2013. DOI: [10.1016/j.wocn.2013.07.002](#).
- [16] Q. Hu, K. Richmond, J. Yamagishi, and J. Latorre, “An experimental comparison of multiple vocoder types,” in *8th ISCA Workshop on Speech Synthesis*, Barcelona, Spain, Aug. 2013, pp. 155–160. [\[pdf\]](#)
- [17] Z. Ling, K. Richmond, and J. Yamagishi, “Articulatory control of HMM-based parametric speech synthesis using feature-space-switched multiple regression,” *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 21, no. 1, pp. 207–219, Jan. 2013, ISSN: 1558-7916. DOI: [10.1109/TASL.2012.2215600](#).
- [18] K. Richmond, Z. Ling, J. Yamagishi, and B. Ura, “On the evaluation of inversion mapping performance in the acoustic domain,” in *Proc. Interspeech*, Lyon, France, Aug. 2013. [\[pdf\]](#)
- [19] J. Scobbie, A. Turk, C. Geng, S. King, R. Lickley, and K. Richmond, “The Edinburgh speech production facility DoubleTalk corpus,” in *Proc. Interspeech*, Lyon, France, Aug. 2013. [\[pdf\]](#)
- [20] I. Steiner, K. Richmond, and S. Ouni, “Speech animation using electromagnetic articulography as motion capture data,” in *Proc. 12th International Conference on Auditory-Visual Speech Processing*, 2013, pp. 55–60. [\[pdf\]](#)
- [21] Z.-H. Ling, K. Richmond, and J. Yamagishi, “Vowel creation by articulatory control in HMM-based parametric speech synthesis,” in *Proc. Interspeech*, Portland, Oregon, USA, Sep. 2012. [\[pdf\]](#)
- [22] Z. Ling, K. Richmond, and J. Yamagishi, “Vowel creation by articulatory control in HMM-based parametric speech synthesis,” in *Proc. The Listening Talker Workshop*, Edinburgh, UK, May 2012, p. 72. [\[pdf\]](#)
- [23] K. Richmond and S. Renals, “Ultrax: an animated midsagittal vocal tract display for speech therapy,” in *Proc. Interspeech*, Portland, Oregon, USA, Sep. 2012. [\[pdf\]](#)
- [24] I. Steiner, K. Richmond, I. Marshall, and C. D. Gray, “The magnetic resonance imaging subset of the mngu0 articulatory corpus,” *The Journal of the Acoustical Society of America*, vol. 131, no. 2, EL106–EL111, Jan. 2012. DOI: [10.1121/1.3675459](#).

- [25] I. Steiner, K. Richmond, and S. Ouni, “Using multimodal speech production data to evaluate articulatory animation for audiovisual speech synthesis,” in *3rd International Symposium on Facial Analysis and Animation*, Vienna, Austria, 2012. [\[pdf\]](#)
- [26] B. Uria, I. Murray, S. Renals, and K. Richmond, “Deep architectures for articulatory inversion,” in *Proc. Interspeech*, Portland, Oregon, USA, Sep. 2012. [\[pdf\]](#)
- [27] J. P. Cabral, S. Renals, J. Yamagishi, and K. Richmond, “HMM-based speech synthesiser using the LF-model of the glottal source,” in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, May 2011, pp. 4704–4707. DOI: [10.1109/ICASSP.2011.5947405](#).
- [28] Z.-H. Ling, K. Richmond, and J. Yamagishi, “Feature-space transform tying in unified acoustic-articulatory modelling of articulatory control of HMM-based speech synthesis,” in *Proc. Interspeech*, Florence, Italy, Aug. 2011, pp. 117–120. [\[pdf\]](#)
- [29] L. Ming, J. Yamagishi, K. Richmond, Z.-H. Ling, S. King, and L.-R. Dai, “Formant-controlled HMM-based speech synthesis,” in *Proc. Interspeech*, Florence, Italy, Aug. 2011, pp. 2777–2780. [\[pdf\]](#)
- [30] K. Richmond, P. Hoole, and S. King, “Announcing the electromagnetic articulography (day 1) subset of the mngu0 articulatory corpus,” in *Proc. Interspeech*, Florence, Italy, Aug. 2011, pp. 1505–1508. [\[pdf\]](#)
- [31] B. Uria, S. Renals, and K. Richmond, “A deep neural network for acoustic-articulatory speech inversion,” in *Proc. NIPS 2011 Workshop on Deep Learning and Unsupervised Feature Learning*, Sierra Nevada, Spain, Dec. 2011. [\[pdf\]](#)
- [32] J. Cabral, S. Renals, K. Richmond, and J. Yamagishi, “Transforming voice source parameters in a HMM-based speech synthesiser with glottal post-filtering,” in *Proc. 7th ISCA Speech Synthesis Workshop (SSW7)*, NICT/ATR, Kyoto, Japan, Sep. 2010, pp. 365–370. [\[pdf\]](#)
- [33] D. Felps, C. Geng, M. Berger, K. Richmond, and R. Gutierrez-Osuna, “Relying on critical articulators to estimate vocal tract spectra in an articulatory-acoustic database,” in *Proc. Interspeech*, Sep. 2010, pp. 1990–1993. [\[pdf\]](#)
- [34] G. Hofer and K. Richmond, “Comparison of HMM and TMDN methods for lip synchronisation,” in *Proc. Interspeech*, Makuhari, Japan, Sep. 2010, pp. 454–457. [\[pdf\]](#)
- [35] G. Hofer, K. Richmond, and M. Berger, *Lip synchronization by acoustic inversion*, Poster at Siggraph 2010, Los Angeles, USA, 2010. [\[pdf\]](#)
- [36] Z.-H. Ling, K. Richmond, and J. Yamagishi, “An analysis of HMM-based prediction of articulatory movements,” *Speech Communication*, vol. 52, no. 10, pp. 834–846, Oct. 2010, ISSN: 0167-6393. DOI: [10.1016/j.specom.2010.06.006](#).
- [37] —, “HMM-based text-to-articulatory-movement prediction and analysis of critical articulators,” in *Proc. Interspeech*, Makuhari, Japan, Sep. 2010, pp. 2194–2197. [\[pdf\]](#)

- [38] K. Richmond, R. Clark, and S. Fitt, “On generating Combilex pronunciations via morphological analysis,” in *Proc. Interspeech*, Makuhari, Japan, Sep. 2010, pp. 1974–1977. [\[pdf\]](#)
- [39] A. Turk, J. Scobbie, C. Geng, B. Campbell, C. Dickie, E. Dubourg, E. G. Bard, W. Hardcastle, M. Hartinger, S. King, R. Lickley, C. Macmartin, S. Nakai, S. Renals, K. Richmond, S. Schaeffler, K. White, R. Wiegand, and A. Wrench, *An Edinburgh speech production facility*, Poster presented at the 12th Conference on Laboratory Phonology, Albuquerque, New Mexico. Jul. 2010. [\[pdf\]](#)
- [40] A. Turk, J. Scobbie, C. Geng, C. Macmartin, E. Bard, B. Campbell, C. Dickie, E. Dubourg, B. Hardcastle, P. Hoole, E. Kanaida, R. Lickley, S. Nakai, M. Pouplier, S. King, S. Renals, K. Richmond, S. Schaeffler, R. Wiegand, K. White, and A. Wrench, “The Edinburgh Speech Production Facility’s articulatory corpus of spontaneous dialogue,” *The Journal of the Acoustical Society of America*, vol. 128, no. 4, pp. 2429–2429, 2010. DOI: [10.1121/1.3508679](#).
- [41] J. Cabral, S. Renals, K. Richmond, and J. Yamagishi, “HMM-based speech synthesis with an acoustic glottal source model,” in *Proc. The First Young Researchers Workshop in Speech Technology*, Apr. 2009. [\[pdf\]](#)
- [42] Z. Ling, K. Richmond, J. Yamagishi, and R. Wang, “Integrating articulatory features into HMM-based parametric speech synthesis,” *IEEE Transactions on Audio, Speech and Language Processing*, vol. 17, no. 6, pp. 1171–1185, Aug. 2009, **IEEE SPS 2010 Young Author Best Paper Award**. DOI: [10.1109/TASL.2009.2014796](#).
- [43] K. Richmond, “Preliminary inversion mapping results with a new EMA corpus,” in *Proc. Interspeech*, Brighton, UK, Sep. 2009, pp. 2835–2838. [\[pdf\]](#)
- [44] K. Richmond, R. Clark, and S. Fitt, “Robust LTS rules with the Combilex speech technology lexicon,” in *Proc. Interspeech*, Brighton, UK, Sep. 2009, pp. 1295–1298. [\[pdf\]](#)
- [45] I. Steiner and K. Richmond, “Towards unsupervised articulatory resynthesis of German utterances using EMA data,” in *Proc. Interspeech*, Brighton, UK, Sep. 2009, pp. 2055–2058. [\[pdf\]](#)
- [46] J. Cabral, S. Renals, K. Richmond, and J. Yamagishi, “Glottal spectral separation for parametric speech synthesis,” in *Proc. Interspeech*, Brisbane, Australia, Sep. 2008, pp. 1829–1832. [\[pdf\]](#)
- [47] Z.-H. Ling, K. Richmond, J. Yamagishi, and R.-H. Wang, “Articulatory control of HMM-based parametric speech synthesis driven by phonetic knowledge,” in *Proc. Interspeech*, Brisbane, Australia, Sep. 2008, pp. 573–576. [\[pdf\]](#)
- [48] C. Qin, M. Carreira-Perpinan, K. Richmond, A. Wrench, and S. Renals, “Predicting tongue shapes from a few landmark locations,” in *Proc. Interspeech*, Brisbane, Australia, Sep. 2008, pp. 2306–2309. [\[pdf\]](#)
- [49] I. Steiner and K. Richmond, “Generating gestural timing from EMA data using articulatory resynthesis,” in *Proc. 8th International Seminar on Speech Production*, Strasbourg, France, Dec. 2008.

- [50] J. Cabral, S. Renals, K. Richmond, and J. Yamagishi, “Towards an improved modeling of the glottal source in statistical parametric speech synthesis,” in *Proc. of the 6th ISCA Workshop on Speech Synthesis*, Bonn, Germany, 2007. [\[pdf\]](#)
- [51] R. A. J. Clark, K. Richmond, and S. King, “Multisyn: open-domain unit selection for the Festival speech synthesis system,” *Speech Communication*, vol. 49, no. 4, pp. 317–330, 2007. DOI: [10.1016/j.specom.2007.01.014](#).
- [52] S. King, J. Frankel, K. Livescu, E. McDermott, K. Richmond, and M. Wester, “Speech production knowledge in automatic speech recognition,” *Journal of the Acoustical Society of America*, vol. 121, no. 2, pp. 723–742, Feb. 2007. [\[pdf\]](#)
- [53] K. Richmond, “A multitask learning perspective on acoustic-articulatory inversion,” in *Proc. Interspeech*, Antwerp, Belgium, Aug. 2007. [\[pdf\]](#)
- [54] —, “Trajectory mixture density networks with multiple mixtures for acoustic-articulatory inversion,” in *Advances in Nonlinear Speech Processing, International Conference on Non-Linear Speech Processing, NOLISP 2007*, M. Chetouani, A. Hussain, B. Gas, M. Milgram, and J.-L. Zarader, Eds., ser. Lecture Notes in Computer Science, vol. 4885, Springer-Verlag Berlin Heidelberg, Dec. 2007, pp. 263–272. DOI: [10.1007/978-3-540-77347-4_23](#).
- [55] K. Richmond, V. Strom, R. Clark, J. Yamagishi, and S. Fitt, “Festival multisyn voices for the 2007 blizzard challenge,” in *Proc. Blizzard Challenge Workshop (in Proc. SSW6)*, Bonn, Germany, Aug. 2007. [\[pdf\]](#)
- [56] R. Clark, K. Richmond, V. Strom, and S. King, “Multisyn voices for the Blizzard Challenge 2006,” in *Proc. Blizzard Challenge Workshop (Interspeech Satellite)*, (<http://festvox.org/blizzard/blizzard2006.html>), Pittsburgh, USA, Sep. 2006. [\[pdf\]](#)
- [57] S. Fitt and K. Richmond, “Redundancy and productivity in the speech technology lexicon - can we do better?,” in *Proc. Interspeech 2006*, Sep. 2006. [\[pdf\]](#)
- [58] K. Richmond, “A trajectory mixture density network for the acoustic-articulatory inversion mapping,” in *Proc. Interspeech*, Pittsburgh, USA, Sep. 2006. [\[pdf\]](#)
- [59] R. A. J. Clark, K. Richmond, and S. King, “Multisyn voices from ARCTIC data for the Blizzard challenge,” in *Proc. Interspeech 2005*, Sep. 2005. [\[pdf\]](#)
- [60] G. Hofer, K. Richmond, and R. Clark, “Informed blending of databases for emotional speech synthesis,” in *Proc. Interspeech*, Sep. 2005. [\[pdf\]](#)
- [61] L. Onnis, P. Monaghan, K. Richmond, and N. Chater, “Phonology impacts segmentation in speech processing,” *Journal of Memory and Language*, vol. 53, no. 2, pp. 225–237, 2005. [\[pdf\]](#)
- [62] R. A. J. Clark, K. Richmond, and S. King, “Festival 2 – build your own general purpose unit selection speech synthesiser,” in *Proc. 5th ISCA workshop on speech synthesis*, 2004. [\[pdf\]](#)
- [63] D. Toney, D. Feinberg, and K. Richmond, “Acoustic features for profiling mobile users of conversational interfaces,” in *6th International Symposium on Mobile Human-Computer Interaction - MobileHCI 2004*, S. Brewster and M. Dunlop, Eds., Glasgow, Scotland: Springer, Sep. 2004, pp. 394–398.

- [64] K. Richmond, S. King, and P. Taylor, “Modelling the uncertainty in recovering articulation from acoustics,” *Computer Speech and Language*, vol. 17, pp. 153–172, 2003. [\[pdf\]](#)
- [65] K. Richmond, “Estimating articulatory parameters from the acoustic speech signal,” PhD thesis, The Centre for Speech Technology Research, Edinburgh University, 2002.
- [66] —, “Mixture density networks, human articulatory data and acoustic-to-articulatory inversion of continuous speech,” in *Proc. Workshop on Innovation in Speech Processing*, Institute of Acoustics, Apr. 2001, pp. 259–276.
- [67] J. Frankel, K. Richmond, S. King, and P. Taylor, “An automatic speech recognition system using neural networks and linear dynamic models to recover and model articulatory traces,” in *Proc. ICSLP*, 2000. [\[pdf\]](#)
- [68] S. King, P. Taylor, J. Frankel, and K. Richmond, “Speech recognition via phonetically-featured syllables,” in *PHONUS*, vol. 5, Institute of Phonetics, University of the Saarland, 2000, pp. 15–34. [\[pdf\]](#)
- [69] A. Wrench and K. Richmond, “Continuous speech recognition using articulatory data,” in *Proc. ICSLP 2000*, Beijing, China, 2000. [\[pdf\]](#)
- [70] K. Richmond, “Estimating velum height from acoustics during continuous speech,” in *Proc. Eurospeech*, vol. 1, Budapest, Hungary, 1999, pp. 149–152. [\[pdf\]](#)
- [71] —, “A proposal for the compartmental modelling of stellate cells in the anteroventral cochlear nucleus, using realistic auditory nerve inputs,” Master’s thesis, Centre for Cognitive Science, University of Edinburgh, Sep. 1997.
- [72] K. Richmond, A. Smith, and E. Amitay, “Detecting subject boundaries within text: a language-independent statistical approach,” in *Proc. The Second Conference on Empirical Methods in Natural Language Processing*, Brown University, Providence, USA, Aug. 1997, pp. 47–54. [\[pdf\]](#)