The impact of cross language on acoustic-to-articulatory inversion and its influence on articulatory speech synthesis

Aravind Illa[1], *Aanish Nair*[2], *Prasanta Kumar Ghosh*[1]

[1]Dept. of EE, Indian Institute of Science, Bangalore - 560012
[2]Dept. of ECE, Georgia Institute of Technology, Atlanta, GA - 30332

aravindi@iisc.ac.in, aanishnair@gatech.edu, prasantg@iisc.ac.in

Abstract

Estimating articulatory representations (ARs) from acoustic features is known as acoustic-to-articulatory inversion (AAI). Various factors of input acoustic features impact the performance of AAI. In this work, we investigate the effect of unseen language on the AAI performance in both seen and unseen speaker conditions. We further perform experiments to analyze how these AAI predictions in unseen language and unseen speaker conditions, in turn, impact the articulatory speech synthesis, i.e., articulatory-to-acoustic forward mapping (AAF). We hypothesize that this investigation enables the exploration of alternative approaches to voice conversion across unseen languages using ARs. Experiments are performed on the AAF model trained using English ARs and evaluated on ARs from unseen speakers speaking different native Indian languages, namely, Hindi, Kannada, Telugu, and Tamil. Experiments reveal that, for AAI, there is a drop in performance due to the mismatch in language in both seen and unseen speaker evaluations. For AAF, subjective evaluations reveal that the synthesized speech quality of non-native (mismatched language) speech is comparable with that of English (matched language).

Index Terms: articulatory speech synthesis, acoustic-to-articulatory inversion, articulatory-to-acoustic forward mapping

1. Introduction

Speech production involves movements of various articulators including, lips, tongue, and velum [1]. Estimating the positions of these articulators from speech acoustics is known as acoustic-to-articulatory inversion (AAI) [2]. On the other hand, estimating acoustic representations from articulatory movements is known as articulatory-to-acoustic forward mapping (AAF). Both these systems, AAI and AAF have been shown to benefit various speech technologies. Various factors like language, speaking rate [3], mode of speaking (neutral and whispered speech) [4], and emotional state affect the movement of articulators thereby impacting the acoustic features. The changes in articulatory and acoustic spaces impact the AAI and AAF mappings.

In [5], it is shown that speaker independent AAI (SI-AAI) trained using data from a particular language could predict articulatory features from acoustics in another language but with a drop in performance. However, in these SI-AAI experiments along with language, the speaker is also unseen. Further investigation is needed to factor out speaker characteristics with a speaker dependent AAI (SD-AAI) where there is a mismatch in language while the speaker is seen. On the other hand, in [6], a speaker dependent AAF model trained with a particular language showed a drop in performance when evaluated on an unseen language. However, the experiments in [6] are performed using directly measured articulatory representations (ARs) and are limited to two multilingual speakers. The findings in [5, 6], indicate that both SI-AAI and SD-AAF performances are affected when evaluated on the unseen language. However, the impact on the performance of AAF is not clear when both the speaker and language are unseen. Also, it would be interesting to investigate how the estimated ARs from SI-AAI would influence AAF in comparison with direct ARs, since direct ARs are not readily available. In [7], an investigation on the performance of AAF is performed using estimated and direct ARs in an unseen speaker evaluation. Experimental results with unseen speakers revealed that the synthesized speech from AAF preserves the linguistic information while carrying the voice characteristics of the seen speaker on which the AAF model is trained. Interestingly, it is shown that AAFs trained with estimated ARs using SI-AAI perform better than the direct measured ARs. However, the experiments carried out in [7] focused only on the English language which is used for both training and testing. In this work, we aim to investigate the impact of unseen language evaluation on the performance of AAI and AAF under seen and unseen speaker conditions.

The objectives of this work are as follows: 1) To study the impact on the performance of AAI when evaluated using an unseen language in both seen and unseen speaker setups, 2) Performance comparison between AAFs trained using direct and estimated ARs, under an unseen language and seen speaker setup, 3) To assess the quality of speech synthesized from AAF driven with estimated ARs obtained from SI-AAI when evaluated using a language and speaker unseen to both the AAI and AAF. This study will help us understand the effect of language and speaker factors on ARs estimated using AAI, and its corresponding impact on the quality of speech synthesized using AAF. We believe that this investigation enables us to explore alternative directions to perform voice conversion across different unseen languages using ARs. In voice conversion, the goal is to modify a source speaker's voice to sound as if it is produced by a target speaker. The advantage of the current approach is that it does not require any parallel data from the source and target speakers for training, and only demands acoustic data from the target speaker. The hypothesis is that the articulatory representation from SI-AAI preserves the linguistic message and normalizes speaker characteristics. In this work, the source speakers are unseen to both SI-AAI and AAF while training. We use the target speaker's ARs estimated using SI-AAI to train the AAF models. Experimental results revealed that estimated ARs of source speakers obtained using SI-AAI have been able to drive the target speaker's AAF model, and synthesize speech in the target speaker's voice in a language unknown to the target speaker.

2. Background

Acoustic to Articulatory Inversion: The mapping from acoustics to articulatory movements is known to be complex and non-linear. Neural networks are known to learn complex and non-linear functions well, and it has been shown that recurrent neural network namely bi-directional long short term memory (BLSTM) achieves the state-of-the-art performance in AAI [8, 2]. Typically, these mappings are learned in a speaker dependent manner (SD-AAI), where training and testing data belongs to the same speaker. For a speaker-independent AAI model (SI-AAI), in general, data from multiple speakers are pooled to train the model [2]. When data from a large number of speakers are used for training, these SI-AAI models are shown to generalize well and are able to predict articulatory movements for an unseen speaker. We deploy BLSTM networks for both SD-AAI and SI-AAI, where the initial layers are BLSTM layers followed by a linear regression layer.

Articulatory-to-Acoustic Forward mapping: The AAF mapping function estimates the acoustic representations from the articulatory movements. Several statistical techniques have been proposed in literature for modeling the AAF mapping, including Gaussian mixture models [9], Hidden Markov models (HMM) [10], and neural network based models [11]. A comparison of their performance has shown that the BLSTM performs the best among all these techniques [12]. Hence, similar to AAI, we use BLSTM network to predict acoustic features (Mel-cepstrum) from ARs.

Proposed Approach: In this work, the preliminary investigation is to determine the impact of language mismatch on the performance of AAI and AAF. The primary interest is to assess the quality of speech synthesized using an AAF mapping whose inputs are estimated ARs obtained from an AAI under unseen language and unseen speaker conditions. To carry out the experiments, we propose the following approach as shown in Fig. 1.

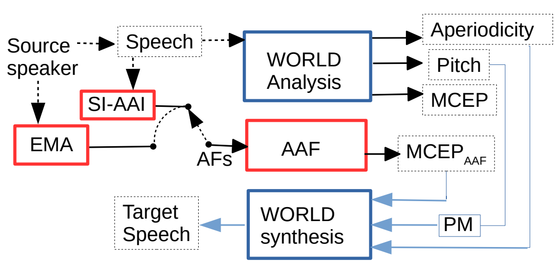

Figure 1: *Illustration of articulatory-to-acoustic forward mapping setup based on the proposed approach.*

We first train an SI-AAI model with multiple reference speakers' acoustic-articulatory data. This SI-AAI model could predict articulatory movements (ARs) for an unseen speaker. An AAF model is trained in a speaker-specific manner (from target speakers list) for whom we are interested to synthesize the voice in a language unknown to the target speakers. The ARs for target speakers are either predicted using the SI-AAI model or measured directly using EMA. As an input to the AAF model, we provide ARs from the target speakers, to estimate acoustic representations in target speaker space. Both the SI-AAI model and the AAF model are trained on English data. During evaluation, a cross speaker's (source speaker unseen for both SI-AAI model and AAF model) speech acoustics are utilized to synthesize speech in the target speaker's voice via the articulatory domain. Speech acoustic features from the cross speaker are fed to the SI-AAI to estimate ARs. The estimated ARs are fed to the AAF model to obtain the acoustic representations. For speech synthesis, we utilize WORLD vocoder [13] to extract pitch and aperiodicity from the cross speaker's speech. We transform original pitch from a source to the target speaker statistics using a linear function [7], indicated as pitch modification (PM) in Fig. 1. The estimated acoustic features from AAF, transformed pitch and aperiodicity are passed through the WORLD synthesizer to synthesize the target speaker's speech.

3. Data collection

For this work, we recorded acoustic-articulatory data using Electromagnetic articulograph AG501 [14]. The EMA recording procedure, speaker details, and speech stimuli used for this work are described below.

EMA recording procedure and post-processing: The EMA recording set-up captures synchronous acoustic and articulatory movements. A t.bone EM9600 shotgun [15], unidirectional electret condenser microphone was used to record the speech data. Articulatory movements were captured using EMA [14]. We recorded six articulatory movements, namely, upper lip (UL), lower lip (LL), jaw (Jaw), tongue tip (TT), tongue body (TB), and tongue dorsum (TD). For head movement correction, we also used two sensors behind the ears [16]. The sensors on the articulators were glued following the guidelines provided in [17]. We considered articulatory movements in the horizontal (X) and vertical (Y) directions in the midsagittal plane, which resulted in a 12-dimensional ARs, which are indicated by UL_x, UL_y, LL_x, LL_y, Jaw_x, Jaw_y, TT_x, TT_y, TB_x, TB_y, TD_x, TD_y.

The recorded acoustic-articulatory data was further post-processed. It is known that the energy of the articulatory trajectories primarily lies below 25Hz, and the movements are slowly varying in nature [18]. Hence, we low-pass filtered the articulatory trajectories at 25Hz to avoid high-frequency noise incurred because of EMA measurement error. The articulatory data was down-sampled from 250Hz to 100Hz. Further, we performed mean and variance normalization for every utterance to remove the effect of sensor average position change across recording and morphological variations across speakers. On the other hand, the acoustics of speech recordings were down-sampled from 48kHz to 16kHz. As an acoustic feature, for AAI experiments, we computed Mel-Frequency Cepstral Coefficients (MFCC) [19], for every 20ms with a shift of 10ms. For AAF models, we decomposed speech signal into their spectral envelope, pitch and aperiodicity using WORLD vocoder [13], and from the spectral envelope we performed mel-scale frequency wrapping and the discrete cosine transform to obtain 36-dim mel-cepstrum (MCEP).

Speech stimuli and Speakers details: For recording with English (EN), we have chosen 460 phonetically balanced English sentences from the MOCHA-TIMIT corpus [20] as the speech stimuli. While for recordings with Indian native languages (NL), namely, Hindi (HN), Kannada (KA), Tamil (TA) and Telugu (TE), we selected 1000 sentences from the text corpus from each language. From these 1000 sentences, we selected a subset of 300 sentences[1] which have phonetically rich coverage using

[1]The sentences in languages other than English were lengthy. In order to keep the duration of the recordings in English and other languages in a similar range, we chose 300 for other languages.

Festivox toolkit [21]. Festivox utilizes greedy selection techniques to select sentences and ensure that the maximum number of different word-word pairings in the corpus are covered by maximizing the bi-gram coverage.

We recorded acoustic-articulatory data from 37 speakers for this work in an age group of 20-28 years. We recorded 460 English sentences from 29 speakers, out of which 25 speakers were used as reference speakers (M1–M13 and F1–F12) to train SI-AAI model and 4 speakers as the target speakers (TM1, TM2, TF1, TF2). From the remaining 8 speakers, we recorded stimuli in native Indian languages with one male and one female in each of the four languages. We refer to them by CF-HI, CM-HI, CF-KA, CM-KA, CF-TA, CM-TA, CF-TE, and CM-TE, where 'C' indicates cross speaker, F/M for gender, HI/KA/TA/TE for language. We also recorded English sentences from all cross speakers except for CM-TE. The following acronyms are used in the paper with respect to AAF and AAI models. SI-AAI – Speaker Independent AAI, SD-AAI – Speaker Dependent AAI, AAF – Speaker Dependent AAF. Note that, we have considered 4 target speakers whose native Indian languages are Bengali and Malayalam. For the unseen condition, we train AAF models using target speakers' English utterances, and cross speakers (utterances from Hindi, Kannada, Tamil, and Telugu) were used as test speakers to evaluate AAF models. This ensures that the target speaker's native language is different from those of cross speakers' native languages.

4. Experimental setup

From all the speakers and languages, 80% of the recorded acoustic-articulatory data is used for training, 10% for validation and 10% for testing.

In experiments with AAI, we investigate the impact of unseen language on AAI in both seen and unseen speaker conditions. In seen speaker condition, we train SD-AAI with cross speakers in their native and English languages, separately. For unseen speaker condition, we train SI-AAI model with reference subjects in English language, and evaluate on cross speakers' data. For both SD-AAI and SI-AAI models, we choose the first three layers as BLSTM layers with 256 units followed by a linear regression output layer with 12 units to predict ARs.

We perform experiments with AAF in seen and unseen speaker conditions. In seen speaker experiments, language impact on AAF is investigated with respect to estimated (using SI-AAI) vs direct ARs (measured with EMA). So, we choose cross speakers and train language dependent AAF models with ARs measured directly and estimated using SI-AAI, separately. In the unseen speaker condition, we train AAF models using target speakers with estimated ARs from SI-AAI model. For AAF model, 12-dimensional ARs are fed as an input and we chose the first three layers as BLSTM layers with 256 units followed by a linear regression output layer.

For evaluating AAI models, we use the Pearson correlation coefficient (CC) as an evaluation metric [18, 2]. In matched case evaluation of AAF, where training and testing is performed with the same target speaker, we use Mel-cepstral distortion (MCD) [7] as an objective measure, which is computed between the original MCEP and predicted MCEP. For unseen speaker and language evaluations, listening tests are carried out.

Listening tests are performed to assess the performance of the proposed approach in terms of (1) naturalness (2) voice similarity and (3) content consistency. The listening tests are conducted with synthesized audio files of 4 Target Speakers using 2 Cross Speakers (Pairs of 1 Male and 1 Female) from the 4 native languages (HN, KA, TA and TE), and English. The tests are conducted in an intra-gender manner i.e., Male to Male and Female to Female. Four listeners volunteered to take part in the test for each native language. For each listener, we present 28 audio files to assess the naturalness comprising 16 files from the target speakers' native language (4 samples x 4 target speakers), 8 files in English (2x4) and 4 files repeated for consistency check. To judge for voice similarity, the listeners listen through 28 pairs of audio files. Each pair is presented with a reference audio file of the target speaker and asked to respond to "are the voices in the two audios from the same speaker?". Finally, to assess content consistency, the listeners listen through 28 pairs of audio files. Each pair is presented with a reference audio file of the cross speaker and the listener is asked to respond to "are the two speakers saying the same words?". In all the three listening tests, listeners rate the audio files on a 5-point scale (5: excellent, 4: good, 3: fair, 2: poor, 1: bad). And all the listening tests were conducted using a web interface.

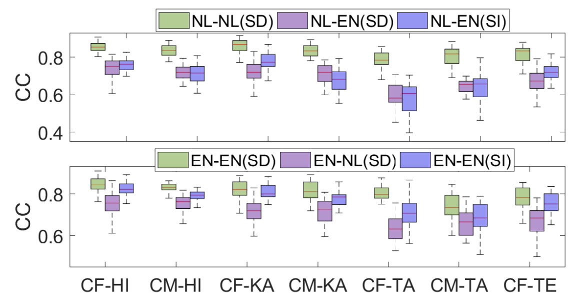

Figure 2: *CC averaged across all articulators for each cross speaker (Evaluation of test sentences from native and English in the top and bottom rows, respectively. In legends "NL/EN" indicates test set language, followed by "NL/EN(SD/SI)" represents the language used to train SD-AAI or SI-AAI model.)*

5. Results and Discussion

In this section, we first present the impact of unseen language on AAI and then on AAF.

5.1. Impact of unseen language on AAI:

In Fig. 2, the top row box-plot reports the results for the native languages (HI, KA, TA and TE) in terms of average CC values for each cross speakers, where the three boxes indicate SD-AAI (matched language), SD-AAI (mismatched language) and SI-AAI (mismatched language and speaker), respectively. Similarly, the second row indicates evaluations on English language. While comparing SD-AAI results in matched and mismatched conditions (NL-NL(SD) vs NL-EN(SD)), there is a drop in performance in NL-EN(SD) compared to NL-NL(SD). The same holds for the English test set. This indicates that in the SD-AAI models, there is a significant drop ($\rho < 0.05$)[2] in AAI performance with language mismatch even though the speaker is seen. Also, we observe that with respect to the matched speaker and language evaluation (EN-EN(SD)/NL-NL(SD)) drop in performance with the unseen speaker is large if there is a mismatch in test-language (NL-EN(SI)) when compared with the

[2]all statistical tests in this work are done by t-test [22]

Table 1: *Subjective evaluation results of AAF models trained with ARs from SI-AAI. AAF models are trained with target speakers ARs in English and tested with cross speakers ARs.*

Language	HINDI	KANNADA	TAMIL	TELUGU	AVG(INDIAN)	ENGLISH
Naturalness	3.25 (0.67)	3.29 (0.64)	3.16 (0.75)	3.55 (0.81)	3.12 (0.71)	3.57 (0.6)
Voice similarity	3.62 (0.63)	3.44 (0.68)	3.28 (0.66)	3.52 (0.62)	3.46 (0.64)	3.39 (0.68)
Content consistency	3.73 (0.59)	3.85 (0.68)	3.73 (0.63)	3.76 (0.78)	3.76 (0.67)	3.75 (0.53)

matched case (EN-EN (SI)). To quantify this for cross speakers, we compute the percentage drop in correlation (PDCC) as follows: PDCC$_L$= $\frac{CC_m - CC_*}{CCm} \times 100$, where CC_m indicates the matched speaker and language with SD-AAI, and CC_* indicates the EN/NL language with SI-AAI model. Across all cross speakers, we observe that for native languages PDCC$_{NL}$ is found to be $15.66\%(\pm 4.87\%)$, while for English PDCC$_{EN}$ is found to be $5.02\%(\pm 2.64\%)$. The results with AAI experiments indicate that there is a significant drop ($\rho < 0.05$) in AAI performance when there is a mismatch in language both in seen and unseen speaker condition. Next, we will investigate how this performance drop in the predicted ARs from SI-AAI impacts the speech synthesis quality of AAF models.

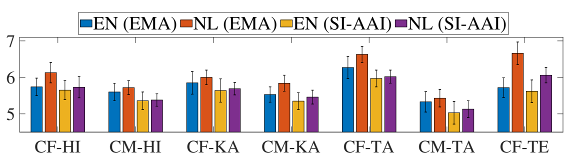

Figure 3: *MCD values of cross speakers obtained by the matched and mismatched language case evaluation on the AAF model trained with English. In legend 'NL/EN' indicates the test set language, followed by '(EMA/SI-AAI)' indicates 'direct ARs' or 'estimated ARs from SI-AAI' used for training and testing AAF.*

5.2. Impact of unseen language on AAF:

We first present the results of AAF trained with cross speakers followed by objective and subjective evaluation of AAF trained with target speakers.

Comparison of direct and estimated ARs: With cross speakers' data from native languages and English, we investigate on comparing direct and estimated ARs, when there is a mismatch in language during training and testing of AAF. Fig. 3, reports evaluation results of cross speakers on AAF model trained with English and tested with NL/EN using direct or estimated ARs. We observe that there is a drop in AAF performance when there is a mismatch in language compared to the matched case in both direct and estimated ARs. Further, we observe that the relative drop between matched and mismatched language evaluation in the case of direct ARs is more (6.4%) as compared to the estimated ARs from SI-AAI (3.4%). These results indicate that although there is a drop in CC while obtaining ARs from SI-AAI due to speaker and language mismatch, interestingly these estimated ARs have less relative drop compared to direct ARs in matched speaker AAF evaluations.

Impact of unseen language on AAF with estimated ARs : In these experiments, we assess the speech quality of the target speaker's AAF, when ARs from cross speakers are used as the input with a language unknown to the target speaker. The subjective results of AAF models with mismatched (both speaker and language) evaluations from cross speakers are reported in Table 1. We observe that the average scores of the listening tests from all native languages (second last column in Table 1) is on par with English subjective evaluations. This could suggest that the synthesized speech from AAF is not degrading due to unseen native languages when compared with English. These results suggest that we can utilize ARs from an unknown speaker to synthetically generate speech in an unknown language for a given target speaker. Sample synthesized files from this work are available online[3].

In this work, we have shown that the AAI performance drops when there is a mismatch in language while evaluation. This further needs phone level or bi-gram level analysis to get more insights. This holds for AAF models as well, our future work will focus on understanding the quality of synthesized speech at the sub-set of phonemes in native Indian languages which are absent in English.

6. Conclusions

In this work, we have performed experiments with 37 speakers' acoustic-articulatory data. Experimental results with AAI models during mismatched language evaluations reveal that there is a drop in performance in both seen and unseen speaker evaluations. However, experiments with AAF trained with cross speakers indicate that ARs estimated using SI-AAI perform better than the directly measured ARs. Also, experiments with AAF trained with target speaker reveal that the synthesis quality of non-native (mismatched) speech is comparable with that of English (matched). This indicates that ARs from AAI in unseen speaker and language can be utilized to drive AAF to synthesize speech for a target speaker in an unknown language.

7. Acknowledgements

We thank all subjects who participated in EMA data collection and listening tests. This work is supported by the Department of Science and Technology (DST), Government of India.

8. References

[1] L. Goldstein and C. A. Fowler, "Articulatory phonology: A phonology for public language use," *Phonetics and phonology in language comprehension and production: Differences and similarities*, pp. 159–207, 2003.

[2] A. Illa and P. K. Ghosh, "Low resource acoustic-to-articulatory inversion using bi-directional long short term memory," *Proc. Interspeech 2018*, pp. 3122–3126, 2018.

[3] ——, "The impact of speaking rate on acoustic-to-articulatory inversion," *Computer Speech & Language*, vol. 59, pp. 75–90, 2020.

[3] https://spire.ee.iisc.ac.in/spire/ICASSP22CrossLangAAF/

[4] A. Illa, Nisha Meenakshi G, and P. K. Ghosh, “A comparative study of acoustic-to-articulatory inversion for neutral and whispered speech,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5075–5079.

[5] G. Sivaraman, C. Espy-Wilson, and M. Wieling, “Analysis of acoustic-to-articulatory speech inversion across different accents and languages,” *Proc. Interspeech 2017*, pp. 974–978, 2017.

[6] B. Cao, A. Wisler, and J. Wang, “Is articulation-to-speech synthesis language independent? a pilot study,” in *International Congress of Phonetic Sciences*. ICPhS, 2013.

[7] A. Illa and P. K. Ghosh, “An investigation on speaker specific articulatory synthesis with speaker independent articulatory inversion,” *Proc. Interspeech 2019*, pp. 121–125, 2019.

[8] P. Liu, Q. Yu, Z. Wu, S. Kang, H. Meng, and L. Cai, “A deep recurrent approach for acoustic-to-articulatory inversion,” in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2015, pp. 4450–4454.

[9] T. Toda, A. W. Black, and K. Tokuda, “Mapping from articulatory movements to vocal tract spectrum with Gaussian mixture model for articulatory speech synthesis,” in *Fifth ISCA Workshop on Speech Synthesis*, 2004, pp. 31–36.

[10] K. Nakamura, T. Toda, Y. Nankaku, and K. Tokuda, “On the use of phonetic information for mapping from articulatory movements to vocal tract spectrum,” in *International Conference on Acoustics Speech and Signal Processing Proceedings*, vol. 1. IEEE, 2006, pp. I–I.

[11] S. Aryal and R. Gutierrez-Osuna, “Data driven articulatory synthesis with deep neural networks,” *Computer Speech & Language*, vol. 36, pp. 260–273, 2016.

[12] Z.-C. Liu, Z.-H. Ling, and L.-R. Dai, “Articulatory-to-acoustic conversion using BLSTM-RNNs with augmented input representation,” *Speech Communication*, vol. 99, pp. 161–172, 2018.

[13] M. Morise, F. Yokomori, and K. Ozawa, “WORLD: a vocoder-based high-quality speech synthesis system for real-time applications,” *IEICE TRANSACTIONS on Information and Systems*, vol. 99, no. 7, pp. 1877–1884, 2016.

[14] “3d electromagnetic articulograph, available online: http://www.articulograph.de/, last accessed:21/10/2018.” [Online]. Available: http://www.articulograph.de/

[15] “EM 9600 shotgun microphone, available online, last accessed:07/04/2020.” [Online]. Available: http://www.tbone-mics.com/en/product/information/details/the-tbone-em-9600-richtrohr-mikrofon/

[16] C. Kroos, “Using sensor orientation information for computational head stabilisation in 3d electromagnetic articulography (EMA),” in *Proceedings of the 10th Annual Conference of the International Speech Communication Association (Interspeech)*. ISCA, 2009, pp. 776–779.

[17] A. K. Pattem, A. Illa, A. Afshan, and P. K. Ghosh, “Optimal sensor placement in electromagnetic articulography recording for speech production study,” *Computer Speech & Language*, vol. 47, pp. 157–174, 2018.

[18] P. K. Ghosh and S. Narayanan, “A generalized smoothness criterion for acoustic-to-articulatory inversion,” *The Journal of the Acoustical Society of America*, vol. 128, no. 4, pp. 2162–2172, 2010.

[19] S. Young and S. Young, “The HTK Hidden Markov Model toolkit: Design and philosophy,” *Entropic Cambridge Research Laboratory Ltd*, vol. 2, pp. 2–44, 1994.

[20] A. Wrench, “MOCHA-TIMIT,” Department of Speech and Language Sciences, Queen Margaret University College, Edinburgh, speech database, 1999. [Online]. Available: http://sls.qmuc.ac.uk

[21] A. Parlikar, S. Sitaram, A. Wilkinson, and A. W. Black, “The Festvox Indic frontend for grapheme to phoneme conversion,” in *WILDRE: Workshop on Indian Language Data-Resources and Evaluation*, 2016.

[22] T. L. Sincich, *Statistics for Engineering and the Sciences*. Chapman and Hall/CRC, 2016.