Vowel variation in a standard context across four major Australian cities

Felicity Cox and Sallyanne Palethorpe

Centre for Language Sciences, Department of Linguistics, Macquarie University, Australia
felicity.cox@mq.edu.au; sallyanne.palethorpe@mq.edu.au

ABSTRACT

Analysis of regional variation in Australian English (AusE) is limited in scope mainly through lack of access to sufficiently diverse speech corpora. Perhaps this helps to explain the long-held belief that regional phonological variation in AusE is quite restricted. Here we present analyses of vowels in a standard phonetic context from word-list data collected for the large-scale AusTalk corpus. The aim is to establish the first published baseline investigation of regional variation for vowels across speakers from four major Australian cities. Twelve monophthongs and 6 diphthongs were examined from 109 male and female speakers under 35 years from Sydney, Melbourne, Adelaide and Perth. Using discrete cosine transform to capture time-varying formant detail, we found evidence for region-specific variation for a small set of vowels including GOAT, NEAR, GOOSE and THOUGHT. The results highlight the need for more detailed analysis of a wider range of phonetic, stylistic, social and regional contexts.

Keywords: Australian English, vowels, regional variation, vowel inherent spectral change, DCT.

1. INTRODUCTION

Australia is the sixth largest country, almost the size of the USA and 32 times greater than the UK so it might be surprising to learn of a long-held belief that regional phonological variation is restricted in Australia [2, 9, 30]. Several reasons have been proposed to explain the apparent regional homogeneity. Bernard [1] and Trudgill [27] suggest linguistic determinism whereby the mix and proportion of the input varieties were so similar in each of the major settlements that the same phonological outcomes ensued. They proposed that high mobility from the time of European settlement until the escalation of both immigration and internal migration during the 1850s' gold rushes contributed to the lack of regional differentiation (see also [30]). The similarity between regional Australian English (AusE) accents has also been considered a consequence of quite recent European settlement, but may also suggest that national identity has a stronger influence than regional affiliation [4]. However, the idea that AusE lacks regionally-specific variation is based on little empirical evidence. The studies conducted have been limited in scope, mainly through the lack of availability of sufficiently diverse speech data. Researchers now have access to AusTalk [6] – a large audiovisual corpus collected using a range of acquisition tasks to sample speech from ~900 individuals across the country. We use AusTalk to provide the foundation for a broader approach to investigating AusE regional vowel variation.

2. BACKGROUND

AusE retains its historical association with Southern British English (SBE) through (among other things) a largely equivalent phonemic vowel inventory described as containing twelve monophthongs (including /eː/ SQUARE), six diphthongs, and schwa [9]. Few region-specific phonological markers have been empirically examined internal to AusE. Mitchell and Delbridge [23] found GOAT /əʉ/ to be the only vowel to display regional variation in their auditory survey of over 7000 school pupils from across the country. They described GOAT from South Australian Independent school girls as having a 'curiously variable glide' ranging from '[ɛ˖ʊ] to [ɛ˖y˖], and from [ɔʊ] to [ɒy]' compared to the more widespread variant of the time [ʌʊ] [23:84] (see also [26]). Bradley [5] also suggests a possible regionally distributed pattern for NEAR /ɪə/ related to vowel dynamicity (see also [13]) with greater monophthongisation in Sydney, New South Wales (NSW) compared to Perth. Western Australia (WA). Another identified regional variant relates to the GOOSE /ʉː/ vowel. Oasa [26], found a more phonetically retracted GOOSE in Adelaide, South Australia (SA) and Melbourne, Victoria (VIC) compared to Sydney, and he identified region-specific offglides for GOOSE in word final (prepausal) position (i.e. a phonetically fronting glide in Adelaide but a retracting glide in Sydney). Horvath and Horvath [19] showed vocalised /l/ as being more common in SA than in NSW, as was the *pool/pull* merger described in [5, 7, 26]. Finally, the most well documented regional phonological effect is the *celery/salary* merger in VIC [8, 21, 22].

Based on these previous analyses, we predict the following differences between the vowels in non-pre-lateral contexts across the major cities:

1. GOAT /əʉ/ is expected to be distinctive in Adelaide based on evidence from [7, 23, 26]. This effect is related to phonetic retraction of the

first element and increased fronting of the second element of the diphthong in Adelaide speakers compared to those from other cities.
2. NEAR /ɪə/ is expected to be more monophthongal in Sydney [5] compared to Perth [13].
3. GOOSE /ʉː/ is expected to be more fronted in Sydney than Adelaide [3, 7, 26] and to show region-specific offglides as described above.

We expect that the most salient differences between cities that are truly the result of a regional effect will apply to both males and females. In addition, the current analysis will use techniques that may capture some dynamic characteristics of vowels and therefore may illuminate previously unreported patterns.

3. METHOD

3.1. Tokens and participants

The present analysis is based on AusTalk [6] which contains data from four scripted and four spontaneous speech tasks. In one scripted task (recorded on three separate occasions with separate randomisations), participants produced 322 words (including the 18 stressed vowels of AusE in the hVd context) upon orthographic presentation on a computer. AusTalk participants under 35 years who had completed all of their primary and high school education in one of four major cities were selected: Sydney (17 males, 17 females), Melbourne (8 males, 17 females), Adelaide (13 males, 12 females), Perth (11 males, 14 females). Insufficient data were available for other major centres to provide adequate power for an extended analysis. Data from the selected speakers included between one and three tokens of each hVd word: 5377 tokens in total. Twelve monophthongs (males: n=1658; females: n=1941) and six diphthongs (males: n=821; females: n= 957) were examined.

3.2. Data Processing and Coding

Data were processed in WebMAUS [20] using an AusE model. Automatically generated textgrids were hand corrected for vowel onset and offset following criteria outlined in [28]. Textgrids were imported to Emu(http://ips-lmu.github.io/EMU.html) for formant checking/correction. Formants were automatically tracked using ESPS/Waves (12th order LPC with a 25 ms raised cosine window and a 5 ms frame shift).

3.3. Vowel Inherent Spectral Change

There is a long history of examining vowels using static target-based methods where the stable portion of the vowel is characterised by the frequencies of the first two or three formants (see [16] and references therein). However, phonemic identity and important phonetic information may be cued by dynamic features as the vowel unfolds in time. Vowel Inherent Spectral Change (VISC) – inherent dynamicity associated with the vowel rather than the context – may be a characteristic of all vowels [12, 14, 24, 25, 31]. In line with [28] and [29], we use the first two discrete cosine transform (DCT) coefficients to encode some of the characteristics of time varying frequency information for each individual hertz-scaled formant trajectory [31]. The zeroth DCT coefficient models the mean of the formant trajectory and the first DCT coefficient models the direction and magnitude of individual formant change, i.e. the slope of the formant as it unfolds in time. Such an approach has been successfully used to model English vowels in [17, 18, 28, 29]. As our data are restricted to a standard phonetic context, DCT coefficients were extracted from formants sampled at 30 equally spaced time-points across the entire vowel.

We acknowledge that consonantal context may affect region-specific vowel production as has been described above for pre-lateral contexts. However, it is also important to begin with a baseline analysis, such as in [10] and [15] who showed that a restricted consonantal context may offer 'important information about the inherent dynamic vowel structure' [15:446]. [11] found that regional variation in vowel production was present in the American English Nationwide Speech Project even in hVd words. To establish a baseline regional comparison of AusE vowels, we similarly present an analysis of hVd data with the caveat that this approach can never reflect the rich diversity of productions that occur in the range of phonetic, prosodic and stylistic contexts but nevertheless can provide a framework against which future fine-grained analyses can be compared.

3.4. Data Analysis

Mixed models were run in R using lmer() separately for males' and females' monophthongs and diphthongs (to avoid the necessity to normalise the data and because the two vowel classes behave differently with respect to dynamicity [14]). In each model either the zeroth or first DCT coefficient (DCT0, DCT1) of F1 or F2 was the dependant variable and the fixed factors were Vowel (12 levels for monophthongs and 6 levels for diphthongs) and City (4 levels). Thus 16 models were run – F1 DCT0, F1 DCT1, F2 DCT0, F2 DCT1 for each sex and for monophthongs and diphthongs separately. Random intercepts were included for Speaker and Repetition. A slope was included for Repetition on Speaker. The anova function and resulting chi-square test of significance was used to compare the difference between competing lme models. In each of the 16

analyses the addition of a random slope did not improve the model fit so was removed thus:
model = lmer(dependent variable ~ Vowel * City + (1|Speaker) + (1 |Repetition))

The lmer analyses indicated a significant Vowel*City interaction χ2 ($p<.001$) for all 16 analyses. The χ2 statistics will only be reported where relevant post-hoc results are discussed. Post-hoc analyses were conducted to illuminate City-based effects with Tukey adjusted significance values.

4. RESULTS

Figures 1 and 2 illustrate the vowel spaces and selected diphthong trajectories for females and males respectively. Post-hoc analyses confirm our first hypothesis that GOAT /əʉ/ shows a strong regional effect. The phonetically retracted first element of the GOAT diphthong giving rise to a raising and considerably fronting glide (see Figures 1 and 2) has been previously described as a regional variant of Adelaide [23, 26]. The present analysis confirms this regional effect and shows that the variant has extended to Perth. The main difference illustrating this effect appears in the F2 DCT1 (Vowel*City: female ($\chi^2(15)$ = 107.72, $p<0.001$; male ($\chi^2(15)$ = 93.59, $p<0.001$). Post-hoc results show that Sydney females' productions differ significantly from Perth and Adelaide (*p*<.001). The same applies to males for Sydney vs Perth (p<.05) with the Sydney vs Adelaide difference showing a strong trend (*p*=.0523). Melbourne also differs significantly from Adelaide (females *p*<.001; males *p*<.05) and Melbourne males differ significantly from Perth males. (p<.05). These results suggest in general terms that productions from the western states (SA and WA) differ from those in the eastern states (VIC and NSW) for this vowel.

Our second prediction of a significant effect for NEAR /ɪə/ dynamicity being reduced in Sydney but greater in Perth had mixed support. For males there was a significant F1 DCT0 effect (Vowel * City (χ^2 (15) = 53.857, $p<0.001$). The mean Perth F1 is lower than the other cities indicating a phonetically raised vowel. Post-hoc results show a difference between Perth and Adelaide (*p*<.01) and a trend for Perth vs Melbourne (*p*=.0568) and Sydney (*p*=.0830). For F2 DCT0 (Vowel * City: female (χ^2 (15) = 50.035, $p<0.001$; male (χ^2 (15) = 64.195, $p<0.001$), Sydney females produced a vowel that was on average significantly more fronted than Adelaide (*p*<.01). Similarly, Sydney males produced a vowel that was on average significantly more fronted than those in Perth (*p*<.05). For F2 DCT1 (χ^2 (15) = 93.59, $p<0.001$), Sydney females differed from Melbourne and Adelaide (*p*<.05), whereas males showed an effect for Melbourne vs Sydney and Adelaide (*p*<.01). F2 DCT1 provides some support for reduced dynamicity for this vowel in Sydney but not compared to Perth as predicted. Further analyses in a greater range of contexts is needed to tease apart the complex relationships pertaining to NEAR.

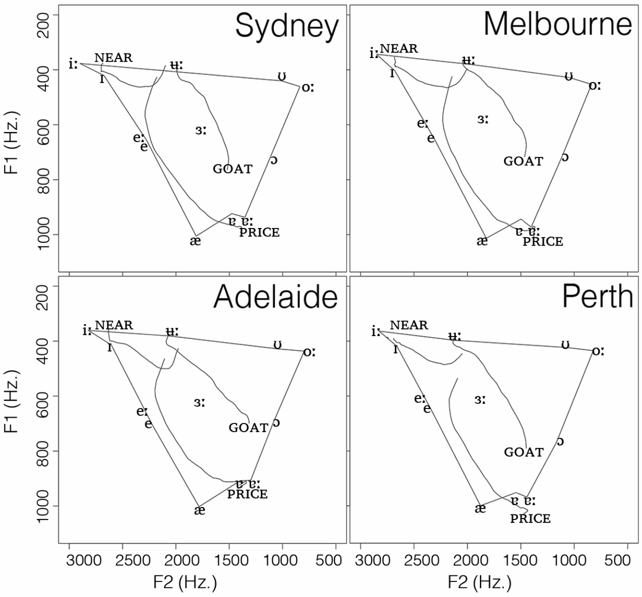

Figure 1: Vowel spaces for female speakers including diphthong trajectories for GOAT /əʉ/, PRICE /ɑe/, NEAR /ɪə/. Note GOAT and PRICE are rising diphthongs and NEAR is a centring diphthong.

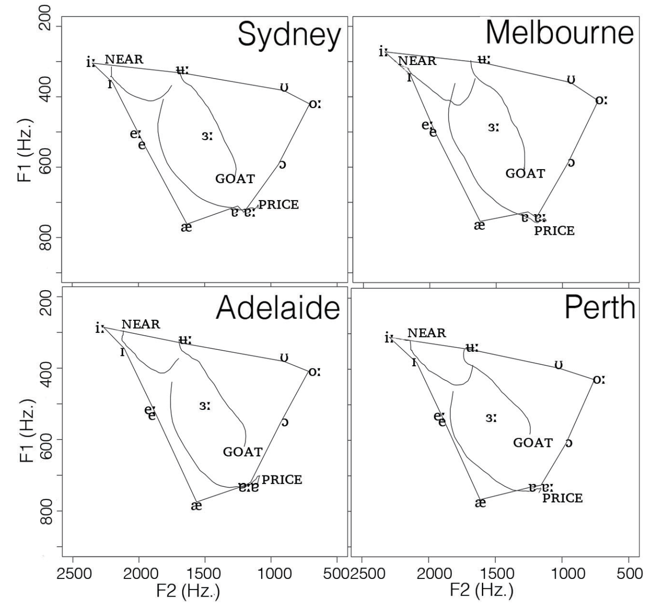

Figure 2: Vowel spaces for male speakers including diphthong trajectories for GOAT /əʉ/, PRICE /ɑe/, NEAR /ɪə/. Note GOAT and PRICE are rising diphthongs and NEAR is a centring diphthong.

The third prediction of a regional effect for GOOSE /ʉː/ was supported but the details require some qualification. The F2 DCT1 results (Vowel * City: female (χ^2 (33) = 141.82, $p<0.001$); male (χ^2 (33) = 145.7, $p<0.001$) show that male and female Adelaide and Perth speakers produce significantly greater fronting as the vowel unfolds than Sydney speakers do as predicted (Sydney vs Adelaide: females $p<.05$; males trend $p=.0527$; Sydney vs Perth: female $p<.001$; male $p<.05$). However, the vowel in Sydney was not found to be generally more fronted as expected (i.e. there was no significant city effect for F2 DCT0). Figure 3 illustrates that Adelaide and Perth pattern together in the F2 trajectory and differ from Sydney.

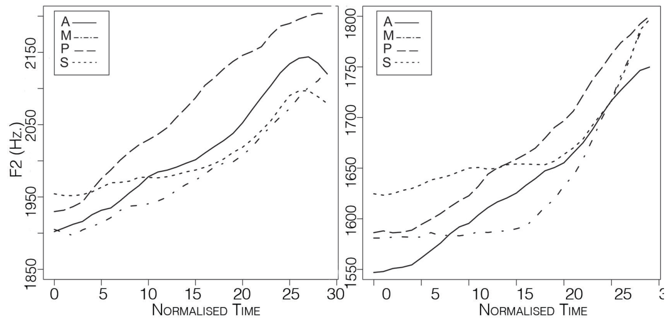

Figure 3: F2 trajectories in hertz over normalised time for GOOSE /ʉː/: females – left panel, males – right panel.

In addition to those stated above, the analysis revealed effects not previously described. Adelaide females produced the THOUGHT /oː/ vowel with greater offglide (progressive raising of F1 indicating phonetic lowering) as exemplified by a significant effect for F1 DCT1 (Vowel * City: female (χ^2 (33) = 61.617, p<0.01). The analysis showed that Adelaide differed from Sydney, Melbourne and Perth ($p<.05$) (Figure 4). Males did not show this difference.

An additional dynamic effect was found for TRAP /æ/ F1 DCT1 (Vowel * City: (χ^2 (33) = 79.095, p<0.0001) showing that Perth males produced less change in F1 across the vowel compared to those from other cities (Perth vs Adelaide $p<.01$, Perth vs Melbourne/Sydney $p<.001$). There was also a significant effect for PRICE F1 DCT0 with Perth females producing phonetically lower vowel (p<.01) than in other cities (Figure 1).

6. DISCUSSION

Figures 1 and 2 show that the monophthong spaces appear highly similar across the four cities and few differences were revealed in our analysis of vowels in this highly controlled phonetic context. Nevertheless, the results of the DCT analysis confirm some of the AusE regional vowel variation suggested in the literature. Our results are consistent with previous reports [7, 23, 26] that GOAT /əʉ/ in SA is characterised by a phonetically retracted onset and a steeply rising and fronting glide. This variant is also produced by Perth speakers (a finding not previously reported) and is present for both males and females. Analysis of NEAR /ɪə/ also shows some regional differentiation but the effects identified were not consistent between males and females and did not show the difference between Sydney and Perth that we had hypothesised [5, 13]. Further investigation of this vowel is warranted. In support of our third hypothesis, GOOSE /ʉː/ displayed diphthongisation (greater gliding of F2) in Adelaide and Perth compared to Sydney [26] but we did not find increased phonetic fronting in Sydney as would be predicted based on analyses in [3, 7, 26].

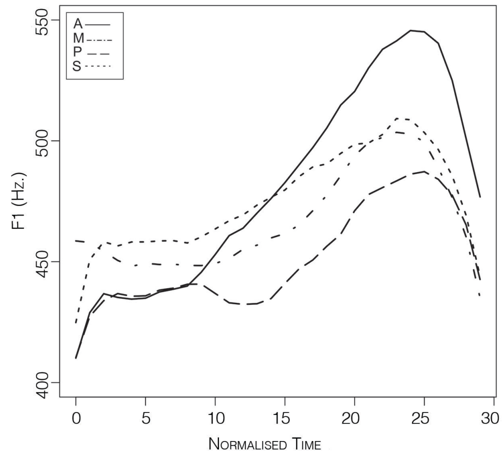

Figure 4: F1 trajectories in hertz over normalised time for female speakers' THOUGHT /oː/.

DCT analysis is a powerful tool in its ability to characterise some aspects of VISC (particularly DCT1). Therefore, it has the potential to highlight effects that would not be revealed using a static formant-based approach. We have found regional differences not previously identified such the offglide of THOUGHT /oː/ used by Adelaide females but not present in other cities. We have also shown, that TRAP for male speakers from Perth has reduced movement through F1 suggesting a flatter trajectory. Intriguingly, [18] also highlight the trajectory of F1 of TRAP in males as differentiating between neighbourhoods in their study of Perth speakers.

7. CONCLUSION

We examined the vowels of AusE in a standard phonetic context produced by speakers under 35 years who had spent all of their schooling in one of the four major Australian cities: Sydney, Melbourne, Adelaide or Perth. The DCT approach has highlighted some regional differentiation in the data and has revealed differences not identified previously in studies based on static targets. This analysis provides a platform for more nuanced examination of contextual effects (in both formal and informal elicitation tasks) that may prove fruitful in further characterising AusE regionally specific variation.

8. REFERENCES

[1] Bernard, J. 1967. *Some Measurements of Some Sounds of Australian English*, Unpublished Doctoral Dissertation, University of Sydney.

[2] Bernard, J. 1981. Australian Pronunciation. In: Delbridge, A. (ed), *The Macquarie Dictionary* (1st edition), 18-27.

[3] Billington, R. 2011. Location, location, location! Regional characteristics and national patterns of change in the vowels of Melbourne Adolescents. *Aust J of Linguistics* 31, 275-303

[4] Blair, D. 1993. Australian English and Australian identity. In: Schultz, G. (ed), *The Languages of Australia*. Canberra: Australian Academy of the Humanities, 62-70.

[5] Bradley, D. 2004. Regional characteristics of Australian English phonology. In: Kortmann, B., Schneider, E., Burridge, K., Mesthrie, R., Upton, C. (eds) *A Handbook of Varieties of English*, Berlin: Mouton de Gruyter, 645-655.

[6] Burnham, D., et al. 2011. Building an audio-visual corpus of Australian English: large corpus collection with an economical portable and replicable Black Box. *Proc. Interspeech* 2011, Florence, 28-31 August 2011.

[7] Butcher, A. 2006. Formant frequencies of /hVd/ vowels in the speech of South Australian females. *Proc. 11th Australian International Conference on Speech Science & Technology*, Auckland.

[8] Cox, F., Palethorpe S. 2004. The border effect: vowel differences across the NSW-Victorian border. *Proc. Australian Linguistic Society Conference*, Newcastle, 2003.

[9] Cox, F., Palethorpe, S. 2007. Illustrations of the IPA: Australian English. *J. Int. Phonet. Assoc.* 37, 341-350.

[10] Clopper, C. G., Pisoni, D.B, de Jong, K. 2005. Acoustic characteristics of the vowel systems of six regional varieties of American English. *J. Acoust. Soc. Am.* 118, 1661- 1676.

[11] Clopper, C. G., Pisoni, D.B. 2004. The Nationwide Speech project: A new corpus of American English dialects. *Speech Comm.* 48, 633-644.

[12] Docherty, G., Foulkes, P., Gonzales, S., Mitchell, N. 2018. Missed connections at the junction of sociolinguistics and speech processing. *Topics in Cognitive Science* 10, 759-774.

[13] Docherty, G., Gonzalez, S., Mitchell, N., Foulkes, P. 2016. The realisation of vowels in conversations by speakers of West AusE. *1st SocioPhonAus Workshop*, Brisbane.

[14] Elvin, J., Williams, D., Escudero, P. 2016. Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English. *J. Acoust. Soc. Am.* 140, 576-581.

[15] Fox, R., Jacewicz, E. 2017. Reconceptualising the vowel space in analyzing regional dialect variation and sound change in American English. *J. Acoust. Soc. Am.* 142, 444-459.

[16] Harrington, J. 2010. *Phonetic Analysis of Speech Corpora.* Chichester: Wiley-Blackwell.

[17] Harrington, J., Kleber, F. Reubold. U. 2008. Compensation for coarticulation, /u/-fronting, and sound change in Standard Southern British: An acoustic and perceptual study. *J. Acoust. Soc. Am* 123. 2825–2835.

[18] Harrington, J., Schiel, F. 2017. /u/-fronting and agent-based modelling: The relationship between the origin and spread of sound change. *Language* 93, 414-445.

[19] Horvath, B., Horvath, R. 2001. A multilocality study of sound change in progress: The case of /l/ vocalization in New Zealand and Australian English. *Language Variation and Change* 13, 37-57.

[20] Kisler, T., Reichel, U. D., Schiel, F. 2017. Multilingual processing of speech via web services. *Computer Speech & Language.* 45, 326–347.

[21] Loakes, D., Clothier, J., Hajek, J., Fletcher, J. 2014. An Investigation of the /el/-/ael/ Merger in Australian English: A Pilot Study on Production and Perception in South-West Victoria. *Aust J of Linguistics* 34, 436-452.

[22] Loakes, D. Hajek, J., Fletcher, J 2017. Can you t[æ]l I'm from M[æ]lbourne? *Eng World-Wide*, 38, 29-49.

[23] Mitchell, A. G., Delbridge, A. 1965. *The Speech of Australian Adolescents*. Sydney: Angus and Robertson.

[24] Morrison, G. S. 2013. Vowel inherent spectral change in forensic voice comparison. In Morrison, G. S., Assmann, P. F. (eds), *Vowel inherent spectral change*. Heidelberg, Germany: Springer-Verlag, 263–283.

[25] Nearey, T., Assmann, P. 1986. Modelling the role of inherent spectral change in vowel identification. *J. Acoust. Soc. Am.* 80, 1297–1308.

[26] Oasa, H. 1989. Phonology of current Adelaide English, In Collins, P., Blair, D. (eds), *Australian English: The Language of a New Society.* St. Lucia: University of Queensland Press, 271-287.

[27] Trudgill, P. 2004. *New Dialect Formation: The Inevitability of Colonial Englishes*, New York: Oxford University Press.

[28] Watson, C., Harrington, J. 1999. Acoustic evidence for dynamic formant trajectories in Australian English vowels. *J. Acoust. Soc. Am.* 106, 458-468.

[29] Williams, D., Escudero, P. 2014. A cross-dialectal acoustic comparison of vowels in Northern and Southern British English. *J. Acoust. Soc. Am.* 136, 2751-2761.

[30] Yallop, C. 2003. A. G. Mitchell and the development of Australian pronunciation. *Aust J of Linguistics* 23, 129–141.

[31] Zahorian, S., Jagharghi, A. 1993. Spectral-shape features versus formants as acoustic correlates for vowels. *J. Acoust. Soc. Am.* 94, 1966-1982.