Acoustic merger between /e/ and /æ/ in Singapore English: insights into stylistic variation and sub-varietal difference

Canaan Zengyu Lan, Olga Maxwell, and Chloé Diskin-Holdaway

The University of Melbourne
zengyul@student.unimelb.edu.au, omaxwell@unimelb.edu.au, chloe.diskinholdaway@unimelb.edu.au

ABSTRACT

This study investigates the extent of the /e/-/æ/ acoustic merger in Singapore English produced by young Chinese Singaporeans in their twenties. We examine the effects of speech style on the acoustic properties of these vowels and the amount of spectral overlap. Participants were recorded reading a wordlist and sentences written in Singapore Standard English (SSE) and Singapore Colloquial English (SCE). Results show some evidence of merger between /e/ and /æ/ on the group level, with the SSE and the SCE sentence styles showing almost complete overlap, but the wordlist showing the least overlap. However, these overlapping patterns are highly gradient when examined individually, with a merger present in the wordlist for some of the speakers. These findings raise questions about the status of Singapore English as a stable or emergent variety and emphasise the need for further fine-grained phonetic research on the vowel systems in New Englishes.

Keywords: Acoustic merger, Singapore English, front vowels, speech production.

1. INTRODUCTION

This study focuses on variability in vowel production in Singapore English (henceforth SgE), which is an umbrella term used to describe English spoken in Singapore. We investigate the extent of acoustic overlap between the DRESS and TRAP vowels (henceforth /e/ and /æ/) across different speech styles in SgE, of which there are at least two syntactically, morphologically and phonologically distinct sub-varieties. Singapore Standard English (henceforth SSE) is generally defined as the standard variety spoken in formal situations and Singapore Colloquial English (henceforth SCE), otherwise known as Singlish, is an informal variety used among friends and family [1]–[3].

Speakers of SgE have been found to demonstrate linguistic unity, homogeneity and confidence in the local standard. This development of "pan-Singaporean" features [4] has been attributed to the rising status of English as the first language and most spoken home language amongst young Singaporeans [5]. Meanwhile, there is an ongoing debate about "standardness", more specifically, the use of SSE in co-existence with SCE or Singlish [6] and the perceived fluidity of boundaries between the two sub-varieties among Singaporeans [3]. Consistent phonetic and phonological inter- and intra-speaker variation [3] further raises questions about whether this is a result of stylistic variation (e.g., [7]) or code-switching between SSE and SCE (e.g., [1]). The front vowels are of particular interest in SgE, given the reported loss of acoustic distinction in SCE [8], and a large degree of variability in vowel realisations [1], [7]–[11].

Previous acoustic investigations of SgE have examined /æ/ alongside /e/ (e.g., [8], [9]). They focus on the /e/-/æ/ acoustic merger, a tendency among most SgE speakers to merge the two vowels, resulting in their phonetic properties being closer to [ɛ]. While a body of work (e.g., [7]–[9]) has explored the realisation of /e/, less is known about the spectral variability of /æ/, addressed in the present paper precisely for this reason.

Studies that have examined the extent of /e/-/æ/ merger in formal speech styles in SgE, such as the citation form or reading tasks, reveal that spectral overlap can be more partial due to the unpredictability in /e/ realisations, where some /e/ tokens merge with /æ/ and the other tokens show raising towards the diphthongal [eɪ] position [12], [13]. As first noted in [13], /e/-raising occurs in the speech of some educated young SgE speakers and only in certain words, such as *bed* and *head*. Further studies that find raising in *next* but not *text* suggest an emerging NEXT-TEXT split in SgE (e.g., [7], [9]). This is in line with theoretical concerns raised by [14] that in some cases of lexically specific phonetic variation, often in contexts of new Englishes, a merger exists more as a synchronic rather than a diachronic phenomenon.

Unlike phonologically conditioned /e/-raising before a voiced velar found across many parts of North America [15], work on SgE (e.g., [9]) indicates that such raising seems rather inconsistent, despite the observation of more frequent /e/-raising before voiced stops (e.g., [7]). A recent examination by [7] comparing older and younger age groups corroborates previous observations that not all young speakers exhibit an /e/-split but that the /e/-/æ/ merger is a strong motivator for /e/-raising among younger speakers. Further, [7]'s observation about stylistic stratification among young Chinese Singaporean participants emphasises the importance of both

speech style and age group as key parameters in the investigation of the acoustics of /e, æ/. Moreover, [7] suggests a negative relationship between /e/-raising and speech formality, with more formal speech styles (i.e., a wordlist versus a reading passage) resulting in less /e/-raising.

Most previous linguistic observations of SgE have assumed correlations between speech styles and sub-varieties, where formal style/reading tasks means the production of SSE, and casual style/interviews means the production of SCE (e.g., [16], [17]). The present preliminary study attempts to disambiguate sub-varieties and styles in SgE. We examine read speech elicited with the help of a wordlist and a set of sentences, further subdivided into sentences written in SSE and SCE (see §2.2 for study design). The production data is based on three sub-sets: **wordlist** (SSE, most formal), **SSE sentence** (SSE, formal but less controlled), and **SCE sentence** (SCE, least formal).

The research questions are as follows:

1. How does speech style (most formal to least formal) influence the acoustic characteristics of /e/ and /æ/ in SgE?
2. How do sub-varietal differences (SSE vs. SCE) contribute to variation in SgE?

2. METHODS

2.1. Participants

The study participants were six Chinese Singaporeans (3M, 3F, gender self-identified; age $\bar{x}$=25, s=2.61) who were either enrolled in or had completed their tertiary degree, with no prior residence in an English-speaking country other than Singapore prior to data collection. All participants self-identified as bilingual speakers with L1 English and varying degrees of fluency in Mandarin. This is except for speaker 005_M whose L1 was Mandarin and L2 English.

2.2. Materials and procedure

The data reported here forms part of a larger project on SgE, investigating 14 monophthongal vowels (including the monophthongised diphthongs GOAT and FACE) across a variety of speech styles, including sociolinguistic interviews, which also elicited participants' demographic information. The analyses presented in this paper are based on a wordlist and a set of sentences. The data collection took place remotely: the participants followed a set of guidelines (how to approach each reading task) shared via WhatsApp and recorded themselves using their smartphones. Participants were instructed to place their smartphones in front of them on a table, complete their recording in one sitting, and read as naturally as possible, allowing for a short pause after each sentence. Additionally, participants were instructed to repeat the entire sentence if they stumbled over a word, without stopping the recording to ensure consistency in volume and quality.

For the wordlist, eight target words were embedded in the carrier phrase "Please say ____ again". For the sentence set, two versions of a semantically similar sentence were provided for participants to read aloud, one written in SCE and the other in SSE, with grammar and lexis modified based on the differences mentioned in the literature (e.g., [3]) and the first author's native speaker intuition. There were eight sentence pairs in total, and participants were briefed about the difference between the two versions before reading the randomised sentences. The words "Standard English" or "Singlish" were written before each sentence to further prompt the intended response (e.g., Standard English: *Have you finished eating? Why are you not having your soup?*; Singlish: *Finish eating? Why you not drinking your soup hah?*). The terms Standard English and Singlish (instead of SSE and SCE) were used as they are more common in Singapore. Both the wordlist and the sentence set involved seven repetitions, and only accentually prominent words were considered for the analysis of the sentence set (Table 1).

	/e/	/æ/
wordlist	bet, bed, het, head	bat, bad, hat, had
SSE sentence	anything, else, seven, ten, very	and, apple, can, has, have, having, Japanese, mad
SCE sentence	better, seven, ten, very	and, apple, can, Japanese, mad
Total	529	781

Table 1: Target words and token numbers for /e/ and /æ/, presented by speech style/sub-variety.

2.2. Data processing and analysis

All data were manually annotated in *Praat* using the default setting [18] and analysed via the *emuR* [19] package in *R* [20]. The F1/F2 estimates were automatically extracted at vowel midpoints using *forest* with the parameters set according to speaker gender. All output data were manually inspected in *Praat*, with any datapoints subject to formant tracker errors corrected or removed. Lobanov 2.0 normalisation [21] was applied, combining the data for 14 monophthongs across all speech styles/sub-varieties and speakers, following the assumptions of a vowel-extrinsic formula.

The Pillai-Bartlett trace, or Pillai score, was calculated with the *tidyverse* [22] package. Pillai scores are one of the most conventionally applied approaches to measuring acoustic overlap [23], [24]

and range from 0 (greater overlap) to 1 (greater distinction). A series of Linear Mixed Modelling (LMM) analyses were performed to determine the relationships between the predictors (e.g., speech style, vowel) and the response variables (i.e., normalised F1 and F2) with the inclusion of random effects using *lme4* [25] and *lmerTest* [26]. The final model included vowel, speech style/sub-variety, and the interaction between vowel and speech style/sub-variety as fixed factors. The random intercepts were speaker and word; factor significance was calculated using the Satterthwaite likelihood t-tests. Tukey post-hoc tests were performed to locate the source of differences using the *emmeans* package [27].

3. RESULTS

3.1. Spectral results across speakers

Fig. 3.1 presents normalised F1/F2 ellipse plots for the vowels /e/ and /æ/. For all three speech styles/sub-varieties, there is a degree of overlap for the two target vowels, with the SSE sentence showing the most variation and overlap in the F1/F2 vowel space. The wordlist has the least amount of overlap due to considerable variability in /e/, whereby half of the /e/ tokens are produced in the same vowel space as /æ/, with /æ/ showing the least variation in the wordlist. The other half of the /e/ tokens in the wordlist show raising and some fronting.

Both /e/ and /æ/ show more variation in the SSE sentence sub-set, with /e/ exhibiting fronting as it raises, compared to tokens in the SCE sentence sub-set. The SCE sub-set indicates greater separation for the two target vowels as /æ/ raises, with a reversal of vowel position in which the centroid for /æ/ is somewhat higher and more forward as compared to /e/.

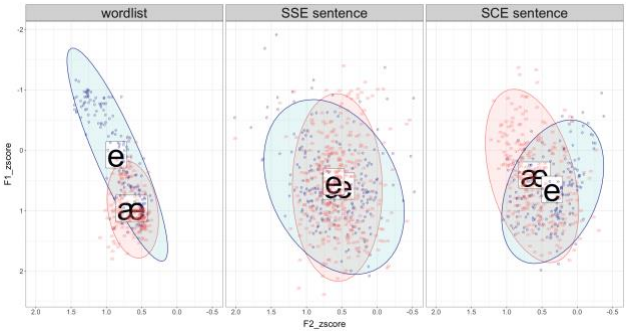

Figure 3.1: Normalised F1/F2 at midpoints of /e/ (blue ellipses), /æ/ (red ellipses) according to speech style/sub-variety (wordlist – left, SSE – middle, SCE – right), with mean F1/F2 estimates of /e/ and /æ/.

The LMM analysis shows no significant effects of vowel or speech style/sub-variety on F1/F2 values. However, the interaction between vowel and speech style/sub-variety is significant for F1 ($p<0.05$), indicating that there are differences between /e/ and /æ/ for the F1 parameter in the wordlist. Further analysis using post-hoc tests reveals a significant difference in F1 ($p<0.05$) and F2 ($p<0.05$) for /e/ between the wordlist and SSE sentence, and in F2 between the wordlist and SCE sentence ($p<0.05$).

In addition, the results revealed potential conditioning effects of the following nasal on vowel formants. The interaction between speech style/sub-variety and the vowel indicates that the F1 of the pre-nasal /æ/ is significantly different from that of the tokens in non-nasal environments in SSE speech style/sub-variety ($p <0.05$). However, caution must be exercised when interpreting these results, given the unbalanced data set (327 out of 1310 tokens where the vowel is followed by a nasal stop). Further investigation using a larger dataset is warranted.

3.2. Individual speaker behaviour

Fig. 3.2 presents the mean F1 (top) and F2 (bottom) estimates for each speech style/sub-variety by speaker. The spectral distributions are highly gradient both individually and stylistically.

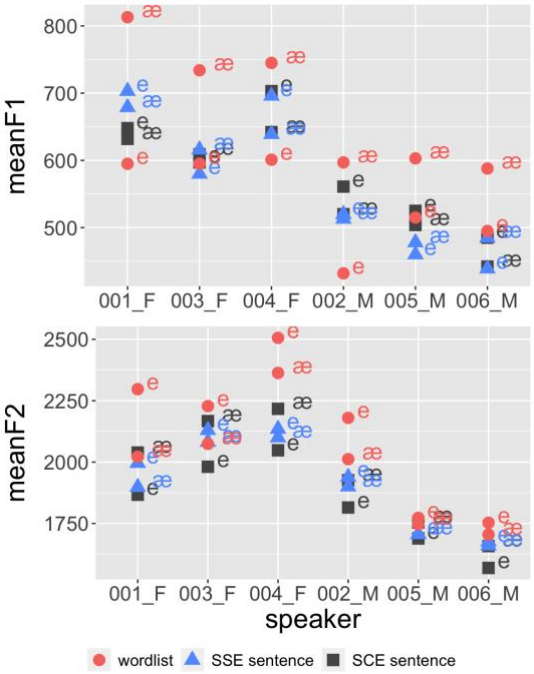

Figure 3.2: Mean F1/F2 estimates (Hertz) (y-axis, F1 – top, F2 – bottom) presented by speaker code in order of gender (x-axis) and speech style/sub-variety (wordlist – red, SSE – blue, SCE – grey)

As shown in the top panel, all participants produce /æ/ lower in the vowel space (i.e., high mean F1 values) in the wordlist (red) in comparison to the sentences (blue, grey). Three speakers (003_F, 005_M, 006_M) raise /e/ in the SSE sentence (blue), i.e., have low F1 values, whereas the other three raise

/e/ in the wordlist (red), at the same time showing a reversal of the /e/ and /æ/ position in the F1/F2 space for the two sentence sets. Among speakers exhibiting this reversal, 001_F produces the two target vowels with varied degrees of distinction in F1, both within and across speech styles/sub-varieties, whereas 004_F, although having an /e/-/æ/ distinction, merges SSE-/æ/ (blue) with SCE-/æ/ (grey), and SSE-/e/ with SCE-/e/. The results for speaker 002_M exhibit difference in F1 in the SCE sub-set, with the mean values for /æ/ clustered with those for SSE-/e, æ/.

As illustrated in the bottom panel, all speakers produce a more fronted /e/ in the wordlist (i.e., the highest mean F2, relative to the speakers' vowel spaces) and a more back target in the SCE sentence (i.e., the lowest F2 mean). For the F2 values, the distribution of /e/ and /æ/ seems to be more consistent for three of the speakers (i.e., 001_F, 003_F, 004_F, all female) showing signs of /æ/ fronting in the SCE sentence. In contrast, male speakers 002_M, 005_M, 006_M produce /e/ and /æ/ with little differences in F2 both within and across speech styles/sub-varieties, except for 002_M whose F2 mean for /e/ is lower in the wordlist.

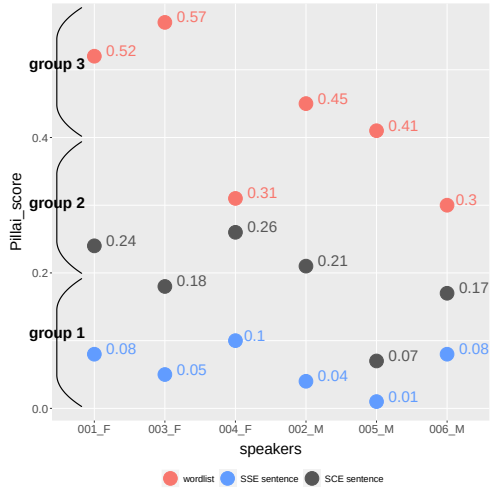

Figure 3.3: Pillai score results with Pillai score on the y-axis and speakers in order of gender on the x-axis, presented by speech style/sub-variety (wordlist – red, SSE – blue, SCE – grey).

Fig. 3.3 illustrates individual Pillai scores. The degree of overlap between /e/-/æ/ across individuals can be grouped according to three main types of overlapping behaviour. Group 1 (scores of 0-0.20, indicating a substantial amount of overlap) reveals the SSE sentence as the most overlapped style/sub-variety for all speakers, with three speakers' (003_F, 005_M, 006_M) SCE sentence also showing a similar degree of overlap. Group 2 (scores of 0.2-0.4, indicating a considerable amount of overlap) includes three speakers (001_F, 004_F, 002_M) whose Pillai scores for the SCE sentence fall within this range, with 004_F and 006_M also having their wordlist showing a similar degree of overlap. Group 3 (scores of 0.4-0.6) includes four speakers (001_F, 003_F, 002_M, 005_M) whose Pillai scores suggest more modest amount of overlap between the two vowels in the wordlist only.

4. DISCUSSION

Returning to the research questions, our analysis has shown that the acoustic results on the group level reveal the least amount of overlap between /e/ and /æ/ for the most formal speech style (i.e., wordlist) and the most overlap for the *less* formal style (SSE sentence), but not the *least* formal style (SCE sentence).

The study confirms [8] and [10]'s findings on /e/-raising in wordlist style in SgE, but not [7]'s findings who reported less frequent /e/-raising in the wordlist. The patterns in our data also exhibit the greatest amount of variation in /e/ and the least variation in /æ/ in the wordlist. We offer two possible explanations. First, hyperarticulation in careful speech could have led to a greater distinction between /e/ from /æ/ in the vowel space. Second, previous research notes that the production of words such as *bed* and *head* is characterised by a higher tendency to be raised by some speakers [13]. This lexical effect was also found in the current study, with different phonetic variants of the close-mid vowel across the target words and individuals, resulting in higher F1 for some of the tokens in the wordlist.

Our analysis also reveals a reversal of /e/ and /æ/ in the SCE sub-set. It is likely that /e/ lowering and backing in the SCE sub-set represent sub-varietal difference between SCE and SSE. Future work will expand the analysis of /e/-/æ/ raising to a much larger dataset and will extend the investigation by applying a dynamic approach, i.e. vowel formant trajectories, including the monophthongised /eɪ/ vowel.

5. CONCLUSION

Our analysis has shown greater /e/-/æ/ overlap in the formal but more controlled style (SSE) as compared to the least formal style (SCE). This pattern potentially reveals complex patterns of variation in SgE, raising questions about acoustic boundaries between speech styles and sub-varieties. The observed /e/-split adds to our knowledge of mergers in (new) Englishes (e.g., [8], [28]), whereby in SgE, a merger may be more accurately conceptualised as a synchronic rather than a diachronic state [14], [29]. This may call for a change of theoretical perspectives and methodological approaches in future studies of mergers and splits in new Englishes.

6. REFERENCES

[1] L. Lim, Ed., *Singapore English: A grammatical description*, vol. G33. Amsterdam: John Benjamins Publishing Company, 2004. doi: 10.1075/veaw.g33.

[2] Z. Bao, *The Making of Vernacular Singapore English: System, Transfer, and Filter*. Cambridge: Cambridge University Press, 2015. doi: 10.1017/CBO9781139135375.

[3] J. R. E. Leimgruber, *Singapore English: Structure, Variation, and Usage*. Cambridge: Cambridge University Press, 2013. doi: 10.1017/CBO9781139225755.

[4] J. H. Sim, "Negotiating social meanings in a plural society: social perceptions of variants of /l/ in Singapore English," Open Science Framework, preprint, Aug. 2021. doi: 10.31219/osf.io/4jbdm.

[5] Census of 2020, "Census of Population 2020 Statistical Release 1: Demographic Characteristics, Education, Language and Religion," Depart of Statistics Singapore, Singapore, 2020. [Online].

[6] D. Y. S. Chye, "Standard English and Singlish: The Clash of Language Values in Contemporary Singapore," presented at the 2009 Conference of the Australian Linguistic Society Conference, 2010, pp. 1–33.

[7] R. L. Starr and A. Choo, "The NEXT TEXT split in Singapore English: Comparing self report and speech production," presented at the 2019 New Ways of Analyzing Variation, University of Oregon, 2019.

[8] D. Deterding, "Emergent patterns in the vowels of Singapore English," *English World-Wide*, vol. 26, no. 2, pp. 179–197, Jun. 2005, doi: 10.1075/eww.26.2.04det.

[9] D. Deterding, "The vowels of the different ethnic groups in Singapore," in *English in Southeast Asia: varieties, literacies and literatures*, D. Prescott, A. Kirkpatrick, H. Azirah, and I. Martin, Eds. Newcastle, UK: Cambridge Scholars Publishing, 2007, pp. 2–29.

[10] B. H. Suzanna and A. Brown, "The [e] and [æ] vowels in Singapore English," in *The English language in Singapore: research on pronunciation*, 2. pr., A. Brown, D. Deterding, and E.-L. Low, Eds. Singapore: Singapore Association for Applied Linguistics, 2002, pp. 84–92.

[11] E. L. Low and A. Brown, *English in Singapore: an introduction*. Singapore ; New York: McGraw Hill, 2005.

[12] D. Deterding, "Phonetics and Phonology," in *Singapore English*, Edinburgh University Press, 2007, pp. 12–39.

[13] M. W. J. Tay, "The Phonology of Educated Singapore English," *English World-Wide*, vol. 3, no. 2, pp. 135–145, Jan. 1982, doi: 10.1075/eww.3.2.02tay.

[14] W. Maguire, L. Clark, and K. Watson, "Introduction: what are mergers and can they be reversed?," *English Language and Linguistics*, vol. 17, no. 2, pp. 229–239, Jul. 2013, doi: 10.1017/S1360674313000014.

[15] J. A. Stanley, "Regional Patterns in Prevelar Raising," *American Speech*, pp. 1–50, Jul. 2021, doi: 10.1215/00031283-9308384.

[16] K. Kalaivanan, F. Sumartono, and Y.-Y. Tan, "The Homogenization of Ethnic Differences in Singapore English? A Consonantal Production Study," *Lang Speech*, vol. 64, no. 1, pp. 123–140, Mar. 2021, doi: 10.1177/0023830920925510.

[17] A. J. Chong, "Towards a model of Singaporean English intonational phonology," presented at the ICA 2013 Montreal, Montreal, Canada, 2013, pp. 060213–060213. doi: 10.1121/1.4800916.

[18] P. Boersma and D. Weenink, "Praat: Doing phonetics by computer." 2022. [Online]. Available: http://www.praat.org

[19] R. Winkelmann, K. Jänsch, S. Cassidy, and J. Harrington, "emuR: Main package of the EMU Speech Database Management System." 2022. [Online]. Available: https://cran.r-project.org

[20] R Core Team, "R: A language and environment for statistical computing." 2022. [Online]. Available: https://www.r-project.org

[21] J. Brand, J. Hay, L. Clark, K. Watson, and M. Sóskuthy, "Systematic co-variation of monophthongs across speakers of New Zealand English," *Journal of Phonetics*, vol. 88, p. 101096, Sep. 2021, doi: 10.1016/j.wocn.2021.101096.

[22] H. Wickham *et al.*, "Welcome to the Tidyverse," *JOSS*, vol. 4, no. 43, p. 1686, Nov. 2019, doi: 10.21105/joss.01686.

[23] J. Nycz and L. Hall-Lew, "Best practices in measuring vowel merger," *The Journal of the Acoustical Society of America*, vol. 134, no. 5, pp. 4198–4198, Nov. 2013, doi: 10.1121/1.4831400.

[24] W. Heeringa and H. Van de Velde, "A New Vowel Normalization for Sociophonetics," in *Interspeech 2021*, Aug. 2021, pp. 4024–4028. doi: 10.21437/Interspeech.2021-1846.

[25] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting Linear Mixed-Effects Models Using lme4," *J. Stat. Soft.*, vol. 67, no. 1, 2015, doi: 10.18637/jss.v067.i01.

[26] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models," *J. Stat. Soft.*, vol. 82, no. 13, 2017, doi: 10.18637/jss.v082.i13.

[27] R. Lenth, "emmeans: Estimated Marginal Means, aka Least-Squares Means." 2022. [Online]. Available: CRAN.R-project.org/package=emmeans

[28] E. W. Schneider, *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press, 2007.

[29] J. Nycz, "New contrast acquisition: methodological issues and theoretical implications," *English Language and Linguistics*, vol. 17, no. 2, pp. 325–357, Jul. 2013, doi: 10.1017/S1360674313000051.