



## Sentence-final particles in Singapore English: Are they pragmatic or phonological?

James Sneed German<sup>1</sup>, Laurent Prévot<sup>1,2</sup>

<sup>1</sup> Aix Marseille Université, CNRS, LPL UMR 7309, 13100, Aix-en-Provence, France

<sup>2</sup> Institut Universitaire de France, Paris, France

james.german@lpl-aix.fr, laurent.prevot@lpl-aix.fr

### Abstract

While the use of sentence-final discourse particles (SFPs) is typically linked to specific interactional or social functions, their realization is also associated with particular positional and intonational requirements. This raises the question of whether the use of SFPs may be partly driven by the phonological characteristics of sentence-final contexts. In [1], we showed that Singapore English *lah* is overrepresented in contexts involving sentence-final stress and underrepresented elsewhere. This is surprising if the use of *lah* is motivated by purely pragmatic considerations, but can be explained if (i) *lah* is recruited where it can relieve tonal crowding, or (ii) *lah* is avoided when it would result in a long sequence of non-prominent syllables. Such behavior is expected to be more prevalent for SFPs (like *lah*) whose pragmatic function is very general, but less prevalent for SFPs with a more specialized function. In this study, we consider the distributions of a wider range of Singapore English SFPs, including *leh*, *lor*, *ah* and *hor*. Overall, these particles were more evenly distributed across prosodic contexts compared to *lah*, suggesting that prosodic context conditions the use of SFPs, but only when this does not interfere with the speaker's intended message.

**Index Terms:** Singapore English, discourse particles, tonal crowding, rhythm rule, stress

### 1. Introduction

The interaction between discourse particles and intonation is particularly interesting given that segmental and supra-segmental aspects of the speech signal are generally assumed to carry different kinds of information. As with the use of other lexical items, the inclusion or exclusion of a given discourse particle is typically viewed as a choice on the part of the speaker based on the content of the message he or she wishes to convey. However, just as many discourse particles have constraints on their syntactic distribution, they can also have intrinsic prosodic characteristics which limit the set of prosodic contexts with which they are compatible. In [1], we showed that the use of the sentence-final particle (SFP) *lah* in Singapore English (SgE) is influenced by the prosodic context present across different sentence-final positions. Specifically, *lah* was more likely than expected to occur after a clause ending in a lexically stressed syllable than after a clause ending in an unstressed syllable. In that study, as here, we assume following [7] and [8] that SgE includes at least two levels of prosodic phrasing, the Accentual Phrase (AP) and the Intonational Phrase (IP), and that f0 contours are most readily explained in terms of abstract tones aligned to the edges of

those phrases. The AP typically consists of a content word plus associated function words to its left, and is marked at its left and right edges by an L and H tone, respectively. Interpellation between the tones results in an f0 rise towards the end of a phrase, which together with AP-final lengthening [21] gives rise to the impression that SgE has word-final stress. The IP consists of one or more APs and carries an additional tone at its right edge. A typical declarative is therefore marked by a combined HL aligned to its right edge.

[19], [8] and others have observed a close correspondence between lexical stress in Brit. English and SgE, though the phonetic and phonological manifestations of stress certainly differ for the two varieties. [8] for example, showed experimentally that the overall f0 height of sentence-initial APs is correlated with differences in (BrE) lexical stress. Additionally, it can be noted that the f0 peak of the utterance-final AP is typically aligned to the lexically stressed syllable (see *compromise*, Fig. 1). Noting that *lah* is intrinsically non-prominent and tends to act as a 'carrier' of the local intonational contour of the utterance it is attached to [2], [3], [4], [5] (see Fig. 2), we considered two possible explanations for the effects found in [1]:

(i) Speakers may avoid using *lah* when doing so would result in a long sequence of unstressed syllables. In other words, principles of eurhythmy [6] predict that long non-prominent sequences are dispreferred. Since *lah* is non-prominent, adding it to the end of a sequence of other non-prominent syllables should also be dispreferred.

(ii) When the final syllable is stressed, the final HL sequence is constrained to occur on a single syllable. Including non-prominent *lah* can potentially relieve tonal crowding in such cases by providing extra segmental material over which to realize the three tones (Fig. 3).

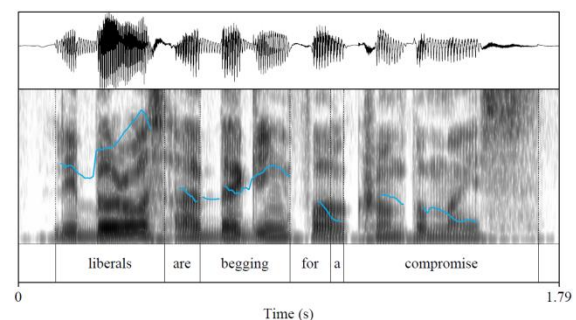


Figure 1. Typical declarative contour of SgE.

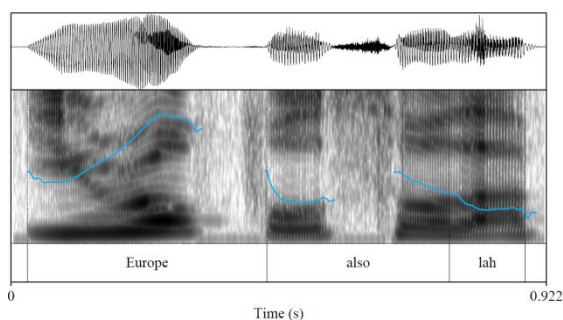


Figure 2: Example of *lah* carrying the local (falling) intonation contour.

If (i) is correct, then the likelihood of *lah* should depend on the specific number of unstressed syllables that occur at the end of an utterance. In other words, as the number of unstressed syllables increases, *lah* should become less likely, since the addition of non-prominent *lah* would degrade such an unstressed string even further. According to (ii), however, the likelihood of *lah* should depend only on whether the final syllable is stressed or not. Our findings showed that the distribution of *lah* did not depend on the specific number of unstressed syllables at the end of the utterance (see Table 1), thus we ruled out (i) in favor of (ii).

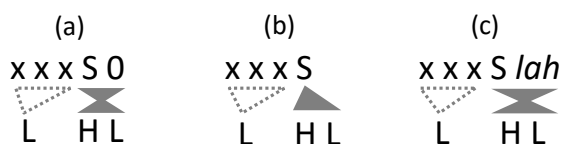


Figure 1. Alignment of utterance-final tones in three prosodic contexts: (a) penultimate stress with no particle, (b) final stress with no particle, and (c) final stress with *lah*.

In that study, we also speculated that the sensitivity of *lah* to prosodic context may be related to the fact that, compared to other SFPs of SgE, *lah* tends to have a rather weak contribution to the meaning of an utterance. Specifically, given that the meaning of *lah* is very general [2], [3], [9], [10], its inclusion or exclusion has relatively little impact on the overall pragmatic force of any given utterance. As long as the meaning that *lah* carries is sufficiently consistent with the speaker's intended meaning, then the speaker may opt to use it even if he or she might not have otherwise. Given that the meaning of *lah* is highly underspecified, this should be the case for a large number of utterances. If our hypothesis is correct, then other SFPs with stronger pragmatic effects than *lah* should show less sensitivity to the prosodic context, since their inclusion would be more likely to interfere with the speaker's intended meaning. In the present study, we test this hypothesis by comparing the distribution of *lah* against four other SFPs, *ah*, *lor*, *hor* and *leh*, which have been shown in a number of studies to make a more substantial contribution to the meaning of the utterances they are associated with.

To give a sense of how the contribution of *lah* can be characterized as 'weak' or 'underspecified', we quote here from one the most comprehensive analyses of SgE particles to date [10: p. 318]:

"... *lah* could be regarded as an indicator to the hearer to proceed with the inferential processes in the derivation of cognitive effects. This characterization of the particle is a rather general one and may seem vague. It is difficult to see how the hearer is expected to gain from the recovery of the explicature. However, there are circumstances where such a signal from the speaker would help her utterance yield a level of relevance consistent with the guarantee communicated by every act of ostensive inferential communication."

A similar idea is echoed in [3: p. 16]:

- (1) A: This is lobster ah?  
B: Lobster lah.

"...in this (constructed) conversation let us assume that A and B are eating lobster, but A isn't sure what it is (could it be prawn?). B wants to assert that it is indeed lobster, so decides to use a particle. If B wants to simply assert that it is lobster, *lah* is appropriate..."

Other related claims involve the notion that *lah* "covers the full range within the assertive continuum" [2: p. 42], that it evokes the meaning "I think that you can know what I want to say" [11: p. 27], that "the range of meanings it possesses is prodigious" [12: 114, as cited in [9]], or that "vagueness may indeed be an inherent property" of *lah* [9: 774]. All of these analyses have in common the idea that the contribution of *lah* to utterance meaning is minimal, and that its use is compatible with a wide range different communicative situations.<sup>1</sup>

This can be contrasted with analyses given to the other SFPs we treat here. *Lor*, for example, has been characterized as marking obviousness [15], directives [2], suggestions [2], or resignation [16], [17]. *Hor* is often linked to the speaker's desire to garner support for a proposition [17] or to soften a directive [10]. [2] classifies *hor* as a 'tentative' particle in contrast with the 'assertive' particles like *lah* or *lor*. *Leh* has been linked to 'maximal assertion', including marking a 'commitment that the interlocutor is expected to act upon' [2: p. 42], as well as to the marking of questions and comparisons [17]. It is generally recognized that *ah* corresponds to at least two different lexical items. One is associated with mid-clausal or inter-clausal breaks and is said to have a 'punctuating' [2] or 'continuation' [17] function. The other is more closely associated with utterance-final positions, and is therefore more relevant to the analysis presented here. It is often characterized as marking the utterance as a question [17] or serving a checking function [2]. According to [2], *lah* and *ah* are associated with a similar range of speech acts and functions, though the author distinguishes them by the fact that *ah* carries a stronger expectation of a response from the interlocutor, whereas *lah* leaves such a response optional. Various authors have argued for more fine-grained distinctions between different variants of these particles. Here we are concerned with more general facts about the overall pragmatic strength of these particles relative to *lah*, so such an analysis is beyond the scope of this study. However, we recognized that such distinctions may eventually lead to more fine-grained predictions in future work, a point which we return to in Section 4.

<sup>1</sup> *Lah* is also commonly characterized as a solidarity marker [12], [13], [14], a function which is arguably compatible with a wide range of communicative situations.

## 2. Methods

For this study, as in [1], we tested our hypothesis through analysis of a text-based corpus of conversational spoken Singapore English (ICE-SIN, [18]). We restricted our analysis to the “Private Dialogues (S1A)” in which the particles of interest are well-represented. Out of a total of 29,855 utterances (comprising 213,555 words) this included 938 tokens of *lah*, 435 tokens of *ah*, 106 tokens of *lor*, 32 tokens of *hor*, and 30 tokens of *leh* in utterance-final positions. We then extracted the last four syllables of each of the 29,855 utterances, not including particles. After excluding words not in our lexical database, this yielded 25,514 utterance contexts with no particle, and 1,435 that included one of the five particles. It should be noted that [2] found *ah* to be the most frequent particle in a corpus of children’s and child-directed speech, whereas here it occurs less than half as often as *lah*. The breakdown across particle types is given in the bottom row of Table 1.

Table 1. Frequency of utterance tokens by prosodic context and particle type.

No. of unstressed syllables	All	<i>lah</i>	<i>ah</i>	<i>lor</i>	<i>hor</i>	<i>leh</i>
0: xxxS	17969	641	247	62	19	23
1: xxS0	7050	198	132	30	4	4
2: xS00	1682	43	17	6	3	0
3: S000	219	6	0	0	0	0
4: 0000	29	0	0	0	0	0
Total	26949	888	396	98	26	27

Overall SgE preserves the stress pattern of British English [19], so utterance-final stress patterns were estimated by cross-indexing wordforms with the Celex2 lexical database [20]. We then categorized utterance tokens according to the number of unstressed syllables intervening between the last stressed syllable in the utterance (primary or secondary) and the end of utterance *excluding* the particle. For example, the utterance “It’s im’possible” would be coded as xS00, while “He’s per’suasive” would be coded as xxS0. The stress patterns of syllables to the left of the last stressed syllable were ignored.

To assess whether the distributions of individual particles are sensitive to the prosodic context, we conducted a series of chi-squared tests. Pearson’s chi-squared can be used to assess whether a frequency distribution conforms to some a priori expectation. In this case, if a particle is not sensitive to prosodic context, then it should be distributed across the different prosodic contexts roughly in proportion to the frequency with which those contexts occur overall. We therefore used the overall frequency distribution as the basis for the expected frequencies in each chi-squared analysis.

## 3. Results

Table 2 gives the proportional frequency distribution for each particle and for all utterances, while Table 3 gives the percentage of deviation from the expected frequency.

Table 2. Proportional frequency of utterance tokens by prosodic context and particle type.

No. of unstressed syllables	All	<i>Lah</i>	<i>ah</i>	<i>lor</i>	<i>hor</i>	<i>leh</i>
0: xxxS	0.667	0.722	0.624	0.633	0.731	0.852
1: xxS0	0.262	0.223	0.333	0.306	0.154	0.148
2: xS00	0.062	0.048	0.043	0.061	0.115	0
3: S000	0.008	0.007	0	0	0	0
4: 0000	0.001	0	0	0	0	0

Table 3. Percentage under- and over-representation of utterance tokens by prosodic context and particle type.

No. of unstressed syllables	<i>lah</i>	<i>Ah</i>	<i>lor</i>	<i>hor</i>	<i>leh</i>	<i>lor, hor, leh</i>
0: xxxS	+8.3%	-6.6%	-5.1%	+9.6%	+27.8%	+3.3%
1: xxS0	-15.2%	+27.6%	+17.0%	-41.2%	-43.6%	-3.8%
2: xS00	-21.0%	-31.1%	-1.9%	+84.5%		-4.4%
3: S000	-17.2%					
4: 0000						

Tables 2 and 3 show that *lah* is somewhat overrepresented in context ‘0’ and underrepresented in other contexts. That this deviates significantly from the overall distribution was confirmed by a chi-square test ( $\chi^2=12.08$ ,  $p < 0.001$ ). By comparison, *ah* appears to be highly overrepresented in context ‘1’ and highly underrepresented in context ‘2’. This was also found to be significant according to a chi-square test ( $\chi^2=15.16$ ,  $p < 0.05$ ). Although the distributions of the other three particles appear to differ from the expected distribution, none of these differences was found to be significant. Since the lack of significance is most likely due to the small sample size, we pooled the distributions of *lor*, *hor* and *leh*. The rightmost column of Table 3 shows that overall, this pooled distribution more closely reflects the overall distribution. Even when the data were pooled in this way, a chi-square test revealed no significant departure from the expected distribution ( $\chi^2 = 1.58$ ,  $p = 0.81$ ).

## 4. Discussion

Overall, our findings show that the particles *ah*, *lor*, *hor* and *leh* do not behave similarly to *lah* with respect to different prosodic contexts. Specifically, in contrast to *lah*, none of these particles is overrepresented when the utterance ends in a stressed syllable. For the pooled distributions of *lor*, *hor*, and *leh* tokens, there was in fact no evidence of a dependence on prosodic context. This is consistent with our hypothesis that *lah* is sometimes recruited to relieve tonal crowding, and that this is made possible by the fact that the inclusion of *lah* is unlikely to conflict with the speaker’s communicative intent. The difference between *lah* and these three particles can be explained by the fact that *lor*, *hor* and *leh* have more substantial consequences for the pragmatic content of an utterance, and are therefore compatible with a much smaller range of situations. In other words, a speaker who seeks to repair a case of tonal crowding is unlikely to recruit one of

these other particles because it would be likely to change the meaning of the utterance in a way the speaker deems undesirable.

While *ah* is not overrepresented in stress-final utterances, it does appear to be sensitive to prosodic context. Specifically, it is highly overrepresented in context ‘1’ and highly underrepresented in context ‘2’. This finding was not expected by any of our earlier hypotheses, and comes as a surprise. Based on tonal crowding or principles of eurhythm, it is difficult to imagine why *ah* appears to have a special relationship to utterances endings in a single unstressed syllable. One possibility is that this effect is lexical – in other words, there is a specific sentence or sentence-final lexical item that *ah* is frequently used with, and this sentence or lexical item happens to end in context ‘1’. Future studies that explore the frequencies of specific lexical contexts are needed in order to determine whether such a phenomenon can explain the direction and size of the observed effects.

In Section 1, we raised the importance of the intrinsic prosodic characteristics of the particles in question. This was important for our analysis of *lah*, since the ability of *lah* to resolve tonal crowding follows from the fact that it is intrinsically non-prominent and can carry the local intonation contour. It may not be the case, however, that all particles have the same intrinsic prosodic characteristics. [17] has suggested that the variant of *ah* most commonly associated with utterance-final contexts in fact has a rising quality (*ah*<sub>24</sub>). Anecdotally, such uses of *ah* involve a rise to the stressed syllable of the final word followed by a fall to the beginning of *ah*, which then rises to the end of the utterance. In the analysis of [7], this corresponds to a LHLH pattern. In case the utterance ends in a stressed syllable, then HLH must be realized on just two syllables. This is in fact the inverse of the case of a stress-final utterance ending in no particle. In that case the presence of the particle solved a problem by contributing extra segmental material while requiring no additional tones (Fig. 3). In the case of *ah*, the particle may actually create a situation of tonal crowding by requiring more tones (two: LH) than the number of extra syllables provided (i.e., one). The difference between the stress-final and penultimate stress contexts with *ah* is depicted in Fig. 4. If this analysis is correct, it could partly explain why *ah* is overrepresented in prosodic contexts involving penultimate stress.

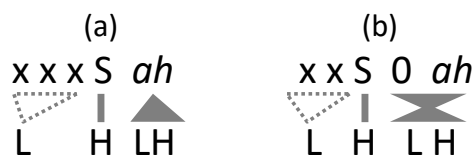


Figure 2. Alignment of utterance-final tones in two prosodic contexts: (a) final stress with ah, (b) penultimate stress with ah.

Establishing whether individual particles have specific tonal requirements will require not only an in-depth analysis of the intonational contours of existing tokens, but also experimental tests to determine the source of these tones. For example, cases like that for *ah* described above may potentially be explained in terms of phrasing requirements (i.e., *ah* projects its own Accentual Phrase), rather than lexical tone, per se. Laboratory experiments might also be useful to

assess the assumption that certain contexts are phonologically degraded. Do listeners accept final and penultimate stress equally well when no particle is present? Does the presence of *lah* improve the acceptability of final stress contexts?

Anecdotally, speakers appear to be able to make adjustments to the prosodic structure of an utterance in order to compensate for changes brought on by the inclusion of a particle. For example, for the utterance “From there originally lah”, which potentially includes an unstressed sequence of three syllables, speakers can shift the major intonational prominence to the final syllable: *original'ly lah*. This type of prominence shifting is in fact apparent in the example in Figure 1, where the major f0 fall begins on the second syllable of *also*, perhaps as a way to maintain the rhythmic alternation of the last three syllables. Thus, laboratory experiments might also be useful to explore speakers do when they are asked to produce potentially degraded contexts.

## 5. Conclusion

In this study, we explored the hypothesis that the sensitivity of the distribution of *lah* to prosodic context could be explained by its pragmatic characteristics. We found some support for this in that certain ‘strong’ particles appear to be insensitive to prosodic context, reflecting the fact that speakers use them in a pragmatically more restricted set of situations. The particle *ah*, however, was highly sensitive to prosodic context, though in a way different from *lah*. Further research will reveal whether this difference can be explained by differences in the intrinsic prosodic characteristics of the two particles.

## 6. Acknowledgements

This research has been made possible through the support of the A\*MIDEX project (n° ANR-11-IDEX-0001-02) funded by the *Investissements d’Avenir* French Government program, managed by the French National Research Agency (ANR). This research was also supported in part by the Erasmus Mundus Action 2 program MULTI of the European Union, grant agreement number 2010-5094-7.

## 7. References

- [1] German, J. S. and Prévot, L. (2014). Phonology constrains the distribution of the particle *lah* in Singapore English. Poster presented at *LabPhon 14*, 25-27 July, Tokyo.
- [2] Gupta, A. R. (1992). The pragmatic particles of Singapore Colloquial English. *Journal of Pragmatics*, 18, 31-57.
- [3] Gupta, A. R. (2006). Epistemic modalities and the discourse particles of Singapore. In Fischer, K. [ed], *Approaches to Discourse Particles*, 243-264. Amsterdam: Elsevier.
- [4] Deterding, D. (1994). The intonation of Singapore English. *Journal of the International Phonetic Association*, 24(02), 61-72.
- [5] Goh, C. (2000). A discourse approach to the description of intonation in Singapore English. In A Brown, A., Deterding, D. and E.-L. Low, *The English language in Singapore: Research on pronunciation*, 35-45. Singapore: Singapore Association for Applied Linguistics.
- [6] Hayes, B. (1984). The phonology of rhythm in English. *Linguistic Inquiry*, 15(1), 33-74.
- [7] Chong, A. (2013). Towards a model of Singaporean English intonational phonology. *Proc. of the Meetings on Acoustics*, 19, Montreal, Canada.
- [8] Chong, A. and German, J. S. (2015). Prosodic phrasing and f0 in Singapore English. *Proceedings of ICPHS 2015*.
- [9] Wong, J. (2004). The particles of Singapore English: A semantic and cultural interpretation. *Journal of Pragmatics*, 36(4), 739-793.
- [10] Ler, V. S.-L. (2005). An in-depth study of discourse particles in Singapore English. Ph.D. dissertation, National University of Singapore.
- [11] Besemeres, M., & Wierzbicka, A. (2003). Pragmatics and cognition: The meaning of the particle *lah* in Singapore English. *Pragmatics & Cognition*, 11(1), 3-38.
- [12] Tongue, R.K. (1974). *The English of Singapore and Malaysia*. Singapore: Eastern Universities Press.
- [13] Richards, J. C. and Tay, M. J. W. (1977). The 'La' particle in Singapore English. In Crewe, W. (ed), *The English language in Singapore*, 141-156. Singapore: Eastern Universities Press.
- [14] Kwan-Terry, Anna. 1978. The meaning and source of the *la* and *what* particles in Singapore English. *RELC Journal* 9(2): 22-36.
- [15] Platt, J., Ho, M.-L. (1989). Discourse particles in Singaporean English: substratum influences and universals. *World Englishes* 2 (8), 215-221.
- [16] Wee, L. (2004). Reduplication and discourse particles. In Lim, L. (ed), *Singapore English: a grammatical description*, 105-126. Amsterdam/Philadelphia: John Benjamins.
- [17] Lim, L. (2007). Mergers and acquisitions: on the ages and origins of Singapore English particles1. *World Englishes*, 26(4), 446-473.
- [18] Greenbaum, S. and Nelson, G. (1996). The international corpus of English project. *World Englishes*, 15(1), 3-15.
- [19] Bao, Z. (2006). Clash avoidance and metrical opacity in Singapore English. *STUF-Sprachtypologie und Universalienforschung*, 59(2/2006), 133-147.
- [20] Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1996). *Celex2. LDC96L14*, Linguistic Data Consortium, Philadelphia.
- [21] Chong, A. J. & German, J. S. (under review). The accentual phrase in Singapore English.