TEMPORAL AND SPECTRAL PROPERTIES OF MADURESE STOPS

Misnadin[1], James P. Kirby[2], Bert Remijsen[3]

Department of Linguistics and English Language, the University of Edinburgh, UK
m.misnadin@sms.ed.ac.uk[1], j.kirby@ed.ac.uk[2], b.remijsen@ed.ac.uk[3]

ABSTRACT

Madurese is a language with a three-way laryngeal contrast and an unusual consonant-vowel co-occurrence restriction. We provide new data on the phonetic realisation of Madurese stops from a sample of 15 native speakers by examining VOT, f0 and two acoustic correlates of voice quality, H1*-H2* and H1*-A3*. Our data indicate that while f0 distinguishes voiced from voiceless (aspirated and unaspirated) stops, at least one voice quality measure contrasts voiced and voiceless aspirated stops with voiceless unaspirated stops, suggesting that the relationship between these features may be more complex than has previously been assumed. Madurese appears to be best described as 'register system' of the Mon-Khmer type, albeit one in which pitch and voice quality are dissociated.

Keywords: Madurese, VOT, f0, register system, voice quality.

1. INTRODUCTION

Madurese is a Western Malayo-Polynesian language spoken primarily on the island of Madura and a number of regions in East Java, Indonesia. One interesting aspect of the language is the fact that it has a three-way voicing distinction in its stop series (voiced, voiceless aspirated and unaspirated) not shared by its neighbouring languages such as Javanese, Indonesian and Sundanese [28].

Previous studies [7-9] examined some acoustic characteristics of the three stop series in Madurese and found that they have significantly different voice onset time (VOT) values. However, VOT values of voiceless aspirated and unaspirated stops were not found to differ markedly [9], which is unexpected in a language with a three-way voicing distinction.

The realization of VOT may be related to another interesting property of Madurese, namely the systematic relation between voicing contrast and vowel height. Specifically, (pre)voiced and voiceless aspirated stops co-occur with high vowels (e.g. [bilis] 'ant', [pʰikʰɤl] 'robber') while voiceless unaspirated stops co-occur with non-high vowels only (e.g. [pɛlɛt] 'massage', [pakaʔ] 'sour') [6-9, 28]. This co-occurrence restriction has been linked to a proposed feature [lowered larynx] ([LL]) by [6, 30]. In addition to explaining the vowel patterning, the feature [LL] is also advanced to account for the reason why [+LL] voiced and voiceless aspirated stops, rather than [-LL] voiceless unaspirated stops, co-occur with high vowels: [LL] spreads rightward until blocked by an existing LL specification, and unspecified vowels become [-LL] be default [6].

As noted by Cohn [6], this hypothesis makes predictions about the phonetic properties of Madurese consonants and the influence they exert on the following vowel. In particular, a lowered larynx predicts that the phonetic realisation of pitch and voice quality of vowels following voiced and aspirated stops may share phonetic properties distinct from those following voiceless unaspirated stops. Preliminary evidence in support of such a difference was observed by [9], who found f0 to be systematically lower after voiced and aspirated stops compared to voiceless unaspirated stops. However, that study involved just two speakers, and as no evidence of voicing during aspirated stops was observed, it remains an open question which, if any, synchronic phonetic properties are shared by voiced and aspirated stops [9].

In this paper, we provide new data on the phonetic realisation of Madurese stops from a larger sample of 15 native speakers. In addition to VOT and f0, we also examine two acoustic correlates of voice quality – H1*-H2* (a measure of open quotient) and H1*-A3* (a measure of spectral tilt) – which are have been successfully used to distinguish voice qualities in a number of languages [12, 17, 19] and which have been mentioned but not examined in previous studies of Madurese [9]. The results will provide phonetic evidence to help us assess the phonologically motivated hypotheses of Cohn [6] and Trigo [30] that Madurese voiced and voiceless aspirated stops share a phonetically grounded phonological feature.

2. METHOD

2.1. Speakers

Fifteen native speakers (8 males, 7 females) of Madurese originating from regencies across Madura

(Bangkalan, Sampang, Pamekasan and Sumenep) were recorded for the study. They were undergraduate students at a university in Madura and all reported no hearing or speech disorders at the time of recording. Their ages ranged from 18 to 28 years. While the participants were also speakers of Indonesian and learnt English, all grew up in Madurese-speaking households and used Madurese in their daily lives.

2.2. Procedures

The stimuli—188 disyllabic Madurese words—were embedded in the sentence frame *Ngereng maos* ____ *se sae* 'Let's read _____ properly' and were presented on a computer screen to each speaker in three random repetitions. They were instructed to read them as naturally as possible. All the recordings were conducted in a quiet room using a Marantz PMD661 portable audio recorder with a Shure SM10A head-mounted microphone. Recordings were divided into three sessions, each of which lasted for approximately 20 minutes. Breaks were provided between sessions.

2.3. Acoustic measurements and analysis

All the segmentations were done manually using Praat [3]. VOT, f0, H1*-H2*, and H1*-A3* were then extracted using available Praat scripts with some modifications when necessary.

It is important to note that H1*-H2* and H1*-A3* denote the corrected measurements for H1-H2 and H1-A3 respectively. That is, H1 and H2 were corrected to undo the effect of F1 while A3 was corrected to eliminate the influence of F1 and F2 [17]. For this purpose, the study employed the improved correction formula proposed by [15] because the formula is applicable to both high and non-high vowels, thus facilitating comparison across vowel types. This correction is particularly important given the covariance between voicing and vowel height in Madurese.

3. RESULTS

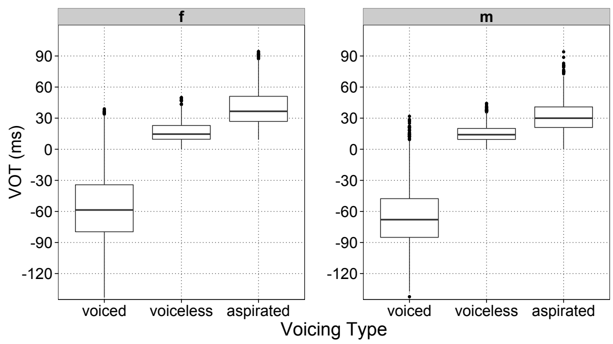
The data were assessed with a series of linear mixed-effects models, which can account for variation due to both random and fixed factors, fit using the *lme4* package [1] for R [25]. The *lmeans* package [26] was used to obtain *p*-values and to perform post-hoc tests. A fixed effect was considered significant at $\alpha = 0.05$. For convenience, each model is specified separately for each dependent variable below.

All of our models included Voicing and Gender along with their interactions as fixed effects, and by-speaker and by-word intercepts for random effects as well as by-speaker random slopes for Voicing. The inclusion of Gender as a fixed effect was motivated by several studies indicating that spectral and temporal properties of stops can vary as a function of speaker gender [16-18, 22, 24, 31].

3.1. VOT

The distribution of VOT is shown in Figure 1. For females, mean VOT for voiceless unaspirated stops is significantly shorter than for voiceless aspirated stops [$\beta = -23.15$, $SE = 2.67$, $t = -8.68$, $p < 0.0001$]. In addition, there is also a significant difference between VOT for voiced and voiceless unaspirated stops [$\beta = 70.77$, $SE = 6.38$, $t = 11.09$, $p < 0.0001$]. For males, VOT differs significantly between voiceless unaspirated and voiced [$\beta = 80.49$, $SE = 6.79$, $t = 11.85$, $p < 0.0001$] as well as voiceless unaspirated and aspirated stops [$\beta = -16.49$, $SE = 2.78$, $t = -5.93$, $p < 0.0001$]. The results also show that females' and males' VOT values for voiceless unaspirated stops and for voiced stops are not significantly different. However, females exhibit significantly longer VOT for voiceless aspirated stops [$\beta = 8.45$, $SE = 3.22$, $t = 2.62$, $p = 0.02$].

Figure 1: Boxplots of VOT as a function of Voicing and Gender (female: left; male: right).

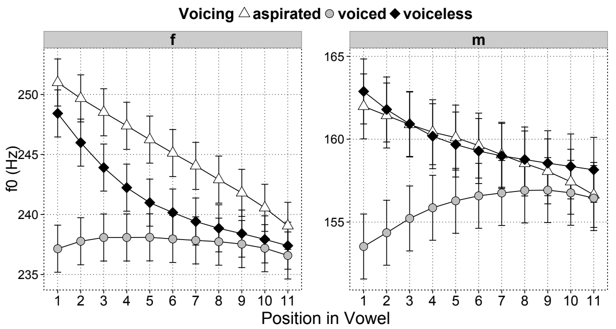


3.2. F0

Mean f0 (averaged over item, speaker and context) for each voicing category is shown in Figure 2. Although f0 was measured at eleven equidistant time-points, here we report statistics for models constructed at just two time-points, vowel onset (point 1) and midpoint (point 6).

For females, the results show that the f0 for voiceless unaspirated stops is significantly different from the f0 for voiced stops at vowel onset [$\beta = 10.88$, $SE = 2.30$, $t = 4.74$, $p < 0.001$], but this difference disappears by vowel midpoint (Fig. 2, left). In contrast, there is no significant difference between the f0 of voiceless unaspirated and aspirated stops at vowel onset, but the difference at vowel midpoint is significant [β = -5.15, *SE* = 1.58, t = -3.25, p < 0.01]. For males, the results indicate that the f0 for voiceless unaspirated stops is significantly different from the f0 of voiced stops at vowel onset [β = 9.46, *SE* = 2.45, t = 3.86, p = 0.003] but not at vowel midpoint. Moreover, no significant difference in f0 values between voiceless aspirated and unaspirated stops was found at either time-point. In summary, the results show that for both genders, the f0 values at vowel onset are higher for voiceless unaspirated and aspirated stops than for voiced stops, while the difference between voiceless aspirated and unaspirated stops is gender-specific.

Figure 2: Mean f0 of vowels measured at 11 equidistant points following voiced, voiceless unaspirated and voiceless aspirated stops (female: left; male: right). Error bars represent 95% confidence interval.



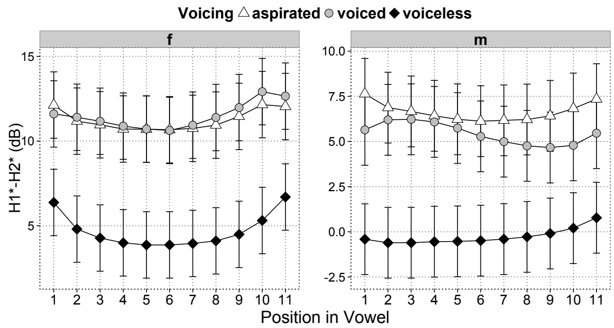
3.3. H1*-H2*

The model specification for H1*-H2* is the same as that for f0 above. Analyses were similarly conducted separately at vowel onset and midpoint.

Mean H1*-H2* (averaged over item, speaker and context) for each voicing category is shown in Figure 3. For females, the results show that H1*-H2* for voiceless unaspirated stops is significantly different from that for voiced stops at vowel onset [β = -5.33, *SE* = 1.63, t = -3.27, p = 0.01]. That is, on average females' H1*-H2* for voiced stops is around 5 dB higher than that for voiceless unaspirated stops. A significant difference was also found for H1*-H2* for voiceless unaspirated and aspirated stops at vowel onset [β = -5.84, *SE* = 1.38, t = -4.23, p = 0.001]. However, there is no significant difference between the H1*-H2* values for voiced and voiceless aspirated stops.

For males, the results show that the H1*-H2* for voiceless unaspirated stops is significantly different from that for voiced stops at vowel onset [β = -5.97, *SE* = 1.74, t = -3.43, p = 0.008]. That is, on average the males' H1*-H2* for voiced stops is around 6 dB higher than that for voiceless unaspirated stops. A significant difference was also found for H1*-H2* for voiceless unaspirated and voiceless aspirated stops [β = -8.02, *SE* = 1.48, t = -5.44, p = 0.0001], but there is no significant difference in H1*-H2* between voiced and voiceless aspirated stops. As seen in Figure 3, a similar pattern also obtains at vowel midpoint for both genders.

Figure 3: Mean H1*-H2* of vowels measured at 11 equidistant points following voiced, voiceless unaspirated and voiceless aspirated stops (female: left; male: right). Error bars represent 95% confidence interval.



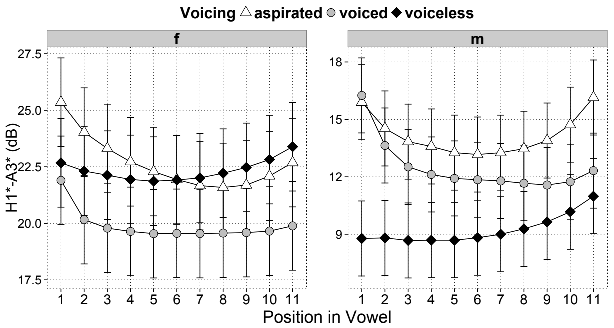
3.4. H1*-A3*

The model specification for H1*-A3* is the same as that for the previous two models and analyses were also performed separately for vowel onset and midpoint.

Mean H1*-A3* (averaged over item, speaker and context) for each voicing category are shown in Figure 4. For females, the results show that H1*-A3* for voiceless unaspirated stops is not significantly different from that for voiced stops at either vowel onset or vowel midpoint. Similarly, H1*-A3* for voiceless unaspirated and aspirated stops do not differ significantly at either time-point. However, H1*-A3* for voiced stops appears to differ significantly from that of voiceless aspirated stops at both vowel onset [β = -3.43, *SE* = 1.14, t = -3.00, p = 0.02] and vowel midpoint [β = -2.40, *SE* = 0.74, t = -3.26, p = 0.009].

For males, H1*-A3* for voiceless unaspirated stops is significantly different from that for voiced stops at vowel onset [β = -7.70, *SE* = 1.67, t = -4.60, p = 0.0006]. A significant difference was also found for the H1*-A3* for voiceless unaspirated and aspirated stops at vowel onset [β = -7.35, *SE* = 1.57, t = -4.67, p = 0.0005], but not for voiced and voiceless aspirated stops. In contrast, no differences in H1*-A3* values between voicing categories were significant at vowel midpoint.

Figure 4: Mean H1*-A3* of vowels measured at 11 equidistant points following voiced, voiceless unaspirated and voiceless aspirated stops (female: left; male: right). Error bars represent 95% confidence interval.



4. DISCUSSION

We found that the three types of Madurese stops have significant VOT differences with voiceless unaspirated and aspirated stops having a relatively small, but significant, difference. This finding is consistent with [8, 9]. Another important finding with regard to VOT in Madurese is the fact that gender-based differences were observed only for voiceless aspirated stops, with females producing slightly longer VOTs than males. It is not clear that physiological factors alone can explain this effect: for instance, while long-lag stops are similarly longer for females than for males in English [29], the opposite effect is observed in Korean [20, 23]. Thus, while gender-based or physiological factors may impact the realization of VOT, it may also depend equally, or more so, on factors such as speech style, social factors, prosodic context, place of articulation, and differences in methodology [19].

We also examined the f0 of vowels following the three stop types and found that for both genders, f0 at vowel onset is higher for voiceless unaspirated and aspirated stops than for voiced stops. Interestingly, however, there is no difference between voiceless unaspirated and aspirated stops at vowel onset. While languages with just a two-way contrast between (phonetically) voiceless aspirated and unaspirated stops tend to show a difference here (e.g. English: [13]; Mandarin: [32]), it is not clear if a difference is to be expected in languages with three-way contrasts. Also interesting is the fact that at vowel midpoint males and females appear to show different f0 patterns. That is, at vowel midpoint, females' f0 values for voiced and voiceless unaspirated stops turn out to be similar, but the voiceless unaspirated and aspirated stops are significantly different. In contrast, males show no differences in their f0 values at vowel midpoint. In general, the present results on f0 in Madurese differ from [9], who observed both voiced and voiceless aspirated stops lower f0, but we suspect this may be due to the fact that there were only two participants in that study.

Finally, we examined two spectral measures of voice quality (H1*-H2* and H1*-A3*) of vowels following each voicing category. Speakers of both genders have consistently higher H1*-H2* values for voiced and voiceless aspirated stops than for voiceless unaspirated stops at both vowel onset and vowel midpoint. While it is true that the vowels following these stop types always agree in height, the correction method employed here has been shown to correct for the effects of F1 on acoustic measures of voice quality [15, 16]. In addition, H1*-H2* values for voiced and voiceless aspirated stops are also consistently similar for both genders at both vowel onset and vowel midpoint.

These results are interesting particularly in relation to the phonological patterning of voiced and voiceless aspirated stops in Madurese. That is, they may shed some phonetic light on the question of why voiced and voiceless aspirated stops pattern together in consonant-vowel interactions in this language. H1*-A3*, on the other hand, provides a less consistent picture, patterning with H1*-H2* for males but not for females. Such differences are in line with [2, 11, 12] who provide evidence that acoustic correlates of voice quality may differ across languages.

In summary, our VOT data are consistent with previous studies on Madurese, but our f0 data are not, instead reflecting the well-established cross-linguistic tendency for f0 to be lower following voiced compared with voiceless stops [18, 24]. Although the H1*-H2* measurement result of the present study is consistent with the proposal of a feature [LL], the fact that f0 does not pattern in this way suggests that the relationship between these features is more complex than has previously been assumed, further underscoring the language-specific nature of the phonetic realisation of laryngeal features [4, 5, 21]. While a perceptual study would be necessary to make any more definitive statements, Madurese appears to be best described as 'register system' of the Mon-Khmer type [e.g. 14], albeit one in which pitch and voice quality appear to dissociate (cf. the decoupling of VOT and voice quality in [27]). Further work on Madurese vowel system is on-going, in order to corroborate previous data showing that [+high] and [-high] vowel sets show systematic differences in F1, but not in F2 [9] as well as to better understand how they interact with the temporal and spectral properties of the stop system presented here.

5. REFERENCES

[1] Bates, D., Maechler, M., Bolker, B., Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-7, <URL: http://CRAN.R-project.org/package=lme4>

[2] Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, *30*, 163-191.

[3] Boersma, P., Weenink, D. (2013). Praat: Doing phonetics by computer [computer program].< www.praat.org>

[4] Chen, Y. (2011). How does phonology guide phonetics in segment-f0 interaction? *Journal of Phonetics, 39*, 612-625.

[5] Chen, Y., Downing, L. (2010). All depressors are not alike. In S. Frota, G. Elordieta, & P. Pilar (Eds.): *Prosodic categories: Production, perception and comprehension: Vol. 82. Studies in natural language and linguistic theory*. Dordrecht: Springer, 243–267.

[6] Cohn, A. C. (1993a). Consonant-vowel interaction in Madurese: The feature lowered larynx. *Chicago Linguistics Society, 29,* 105-119.

[7] Cohn, A. C. (1993b). Voicing and vowel height in Madurese: A preliminary report. *Oceanic Linguistics Special Publications*, *24*, 107–121.

[8] Cohn, A. C., Ham, W. H. (1998). Temporal properties of Madurese consonants: A preliminary report. *Working Papers of the Cornell Phonetics Laboratory*, *12*, 27–51.

[9] Cohn, A. C., Lockwood, K. (1994). A phonetic description of Madurese and its phonological implications. *Working Papers of the Cornell Phonetics Laboratory*, *9*, 67–92.

[10] Esposito, C. M. (2010a). Variation in contrastive phonation in Santa Ana Del Valle Zapotec. *Journal of the International Phonetic Association*, *40*(2), 181-198.

[11] Esposito, C. M. (2010b). The effects of linguistic experience on the perception of phonation. *Journal of Phonetics*, *38*, 306–316.

[12] Gordon, M., Ladefoged, P. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, *29*, 383–406.

[13] House, A. S., Fairbanks, G. 1953. The influence of consonant environment upon the secondary acoustic characteristics of vowels. *Journal of the Acoustical Society of America*, *25*(1), 105–113.

[14] Huffman, F. E. (1976). The register problem in fifteen Mon-Khmer languages. *Austroasiatic Studies*, *Part 1 Oceanic Linguistics Special Publication*(13), 575–589.

[15] Iseli, M., Alwan, A. (2004). An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation," in *Proceedings of ICASSP*, Montreal, Canada, Vol. 1, pp. 669–672.

[16] Iseli, M., Shue, Y., Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *Journal of the Acoustical Society of America, 121(4)*, 2283-2295.

[17] Hanson, H. M. (1997). Glottal characteristics of female speakers Acoustic correlates. *Journal of the Acoustical Society of America*, *101*(1), 466–481.

[18] Hanson, H. M., Chuang, E. S. (1999). Glottal characteristics of male speakers: Acoustic correlates and comparison with female data. *Journal of the Acoustical Society of America 106(2)*, 1064–1077.

[19] Hanson, H. M., Stevens, K. N., Kuo, H.-K. J., Chen, M. Y., Slifka, J. (2001). Towards models of phonation. *Journal of Phonetics*, *29*, 451–480.

[20] Kang, Y. (2014). Voice onset time merger and the development of tonal contrast in Seoul Korean stops: a corpus study. *Journal of Phonetics, 45*, 76-90.

[21] Kingston, J., Diehl, R. L. (1994). Phonetic knowledge. *Language*, *70*(3), 419–454.

[22] Morris, R.J., McRea, C.R., Herring, K.D. (2008). Voice onset time differences between adult males and females: Isolated syllables. *Journal of Phonetics, 36*, 308-317.

[23] Oh, E. (2011). Effects of speaker gender on voice onset time in Korean stops. *Journal of Phonetics, 39*, 59-67.

[24] Peterson, G.E., Barney, H.L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America, 24(2)*, 175-184.

[25] R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

[26] Lenth, R.V., Hervé, M. (2014). lsmeans: Least-Squares Means. R package version 2.13. http://CRAN.R-project.org/package=lsmeans

[27] Pittayaporn, P., Kirby, J. (2014). Laryngeal and tonal contrasts in the Tai dialect of Cao Bang. Poster presented at LabPhon 14, Tokyo.

[28] Stevens, A. M. (1968). *Madurese Phonology and Morphology*. New Haven: American Oriental Society.

[29] Swartz, B.L. (1992). Gender differences in voice onset time. *Perceptual and Motor Skills, 75*, 983-992.

[30] Trigo, L. (1991). On pharynx-larynx interactions. *Phonology*, *8*(1), 113–136.

[31] Whiteside, S. P., Marshall, J. (2001). Developmental trends in voice onset time: Some evidence for sex differences. *Phonetica*, *58*, 196–210.

[32] Xu, C. X., Xu., Y. 2003. Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association, 33* (2): 165–81.