A Preliminary Investigation of the Phonetic Characteristics of Moklen Tones

Warunsiri Pornpottanamas[1], Sireemas Maspong[2] & Pittayawat Pittayaporn[3]

[1]Department of English and Linguistics, Ramkhamhaeng University, Thailand
[2]Institute for Phonetics and Speech Processing (IPS), LMU Munich, Germany
[3]Department of Linguistics & Center of Excellence in Southeast Asian Linguistics, Chulalongkorn University, Thailand

warunsiri.p@rumail.ru.ac.th, s.maspong@phonetik.uni-muenchen.de, pittayawat.p@chula.ac.th

Abstract

Moklen, an endangered Austronesian language in Southern Thailand, is important for studying tonogenesis. Previous research confirmed the presence of two lexical tones in Moklen, but their nature is unclear. This study analyzed the acoustic properties of Moklen tones. Eight native Moklen speakers participated, producing 93 mono- and disyllables with varied tones, onset voicing, vowel length, and coda classes. Acoustic measurements were obtained from the stressed final syllables, including f0, F1, F2, H1*-A3*, and CPP. Results showed that f0 is the primary phonetic cue for tonal contrast in Moklen, accompanied by the difference in vowel quality and phonation type. Specifically, Tone 1 is characterized by higher pitch and a lower and more front vowel with modal voice, while Tone 2 has a lower pitch and a higher, more back, and breathier vowel. These characteristics bear similarities to register distinctions observed in Austroasiatic languages of Southeast Asia, suggesting a possible transphonologization of laryngeal properties into prosodic ones in Moklen. However, the exact segmental sources of Moklen tones still remain an open question.

Index Terms: tones, tonal contrast, phonetic cues, tonogenesis, Moklen, Austronesian language

1. Introduction

Moklen, an endangered Austronesian language spoken on the Andaman coast of Southern Thailand, is an important case for studying tonogenesis [1], [2], [3]. While previous research by Pittayaporn et al. [3] confirmed the presence of two lexical tones in Moklen, the nature of this tonal contrast remains unclear. Swastham [4] and Larish [1], [5], [6] proposed that Moklen tones emerged through contact with Southern Thai, but it is uncertain whether Moklen tones developed from segmental sources following Haudricourt's [7] model of tonogenesis, given that Moklen still maintains contrastive voicing in onsets. For instance, the Proto-Malayo-Polynesian (PMP) *b is preserved in Moklen as /b/, as in *buŋa > /buŋáːʔ/ 'flower' and *bulu > /bulɤ̀j/ 'hair', while the PMP *p has been maintained as /p/, as in *paqit > /pakɛ́t/ 'to be bitter' and *puqun > /pɔkɔ́n/ 'tree'. The two specific issues regarding the phonetic realization of Moklen tones and the underlying mechanisms driving the emergence of these tonal distinctions remain ambiguous. To address the latter question on the development of tones in this language, it is imperative to first uncover the phonetic realization of Moklen tonal contrast. Examining the acoustic characteristics of Moklen tones can provide insights into the hypothesis regarding their origins. In this preliminary study, we conducted an instrumental analysis on the phonetic properties of Moklen tones in different syllable types and onset voicing.

1.1. Moklen language

Moklen, a member of the Austronesian language family, is spoken by approximately 4,000 Moklen individuals residing along the coastline of the southern Thai peninsula. The phonological features of the Moklen language exhibit a remarkable similarity to other mainland Southeast Asian languages, distinguishing it from the insular Austronesian language family [6], [8].

Moklen allows twenty consonants in syllable-initial positions /p, pʰ, b, t, tʰ, d, c, cʰ, k, kʰ, ɡ, ʔ, h, m, n, ɲ, ŋ, l, w, j/, including consonant clusters like /pl/ and /kl/. In the coda of major syllables, ten consonants are possible, including /-p, -t, -k, -ʔ, -h, -m, -n, -ŋ, -w, and -j/.

Additionally, Moklen has nineteen vowels, with nine pairs contrasting for length /i, iː, ɯ, ɯː, u, uː, e, eː, ɤ, ɤːo, oː, ɛ, ɛː, ɔ, ɔː, a, aː/ in the final stressed syllable. In the initial unstressed syllables of disyllabic words, only /a/, /i/, /u/, /ɔ/, and /ɛ/ occur in unstressed syllables. An extra-short neutral vowel /ə/ varies in pronunciation, influenced by its environment. In addition, Moklen includes three diphthongs: /iə/, /ɯə/, and /uə/.

1.2. Moklen tonal phonology

Similar to many languages in Mainland Southeast Asia, Moklen is an iambic language that assigns stress to the last syllable of the foot [1], [3], [4], [5], [6]. While many words are typically disyllabic, they can be simplified especially in connected speech by omitting the initial syllable [1], [3]. The stressed syllable can be either an open or closed syllable with distinct vowels. Disyllabic structures follow the pattern (CV).(C)CV(V)(C), while monosyllables follow (C)CV(V)(C).

Regarding the presence of two lexical tones in Moklen, as confirmed by Pittayaporn et al. [3], tones are consistently realized on the ultimate syllable, which is always stressed. They are categorized as either high or low. Moklen's tonal phonology exhibits an unbalanced distribution, with more words having a high tone than a low one. Notably, only a few tonal minimal pairs have been identified, such as the words /nəmán/ 'to fish,' and /nəmàn/ 'to be glad or happy', the words /niʔúːn/ 'to dry in the sun', and /niʔùːn/ 'coconut', the words /kɔláːt/ 'to be hot', and /kɔlàːt/ 'mushroom', and the words /dulúːk/ 'marlin', and /dulùːk/ 'to light up'. Because it is not possible to predict tones based on surrounding sounds, the language certainly has contrastive tones. These instances prompt us to wonder whether the tonal difference is solely due to pitch, as observed in the tonal Northern and Western dialects of Khmu [9], or involves other acoustic properties, as seen in Eastern Cham [10], Northern Vietnamese [11], Burmese [12], [13].

 10.21437/TAI.2023-13

2. Methodology

2.1. Data collection

To examine the phonetic realization of Moklen tones, eight native Moklen speakers from Phang Nga Province, Thailand, participated in the study. Among them, four were residents of Bang Sak Village (BS), while the others lived in Lam Pi Village (LP). Language data were collected from three females and one male at each research site. The participants ranged in age from 46 to 70 years and were bilingual in both Moklen and Southern Thai, with Moklen being their dominant language.

Participants were instructed to produce Moklen mono- and disyllables in isolation, with each word repeated three times. For the analysis of acoustic properties of Moklen tones, we carefully selected 93 target words having stressed final syllables with /a, aː/ vowels. These target words were systematically varied in terms of tones, onset voicing, vowel length, and coda classes to ensure a balanced representation, as seen in an illustration of stimuli used for eliciting Moklen data in Table 1.

Table 1: *An illustration of Moklen stimuli*.

Onset Voicing	Vowel Length	Coda			
		plain	**-h**	**-q**	**-p, -t, -k**
voice-	short	batáŋ	bəkáh	kapáʔ	baták
less	long	batáːŋ	-	matáːʔ	ʔáːk
					ʔàːk
voiced	short	nəmán	kʰaják	damáʔ	digàt
		nəmàn			namát
	long	bəláː	ʔaláːʔ	namáːʔ	kəláːt
		bəlàː		dadàːʔ	kəlàːt

Note that Moklen displays variation in initial consonants, such as aspirated palatal stops like /cʰ/ sometimes being pronounced as [cʰ] or [s]. Similar variations occur with other consonant pairs like [b] and [m], [d] and [l]. Thus, the word meaning 'stove (traditional)' could be pronounced as [dapán] or [lapán], and the word meaning 'eye' could be pronounced as [matáːʔ] or [batáːʔ]. These patterns of variation appear to be dialectal and generational.

2.2. Data processing

Five acoustic measurements were extracted for analysis: fundamental frequency (f0) for pitch, first and second formant frequencies (F1, F2) for vowel quality, the difference between corrected first harmonics and corrected spectral amplitude of F3 (H1*-A3*)[1], and Cepstral Peak Prominence (CPP) for voice quality. These acoustic measures were chosen, as they are reported to be acoustic correlates of tonation in Southeast Asia [14]. The extraction of acoustic measurements was performed on the vowel intervals of the stressed final syllables, using PraatSauce [15]. A consistent window size of 30 ms with a 5 ms time step was applied across all acoustic measurements.

The acoustic measurement process involved two steps for each participant, following Hirst [16]. In the initial pass, a broad f0 range of 75-400 Hz was used. Subsequently, the first (Q1) and the third quartiles (Q3) of f0 were computed for each participant. In the second round of f0 extraction, the f0 floor was set to $Q1 \times 0.75$, while the f0 ceiling was set to $Q1 \times 1.5$.

2.3. Data analyses

To enable meaningful comparisons, each measurement underwent z-scoring by speakers and time warping to a fixed length. We conducted a third-order polynomial growth curve analysis of f0, F1, F2, H1*-A3*, and CPP trajectories during the vowel interval. The effects of tone, onset voicing, coda class, and their interactions were assessed through hierarchically nested models, following the approach outlined by Mirman [17]. Separate models were conducted for short and long vowels and dialects, resulting in a total of four models for each measure. Subject was incorporated as a random effect in these analyses.

3. Results

Our results show that the acoustic cues for Moklen tonal contrast include pitch, vowel quality and voice quality. In essence, significant differences in f0, F1, F2, H1*-A3*, and CPP were observed between the two tones in both varieties.

3.1. f0

In the case of short vowels, the likelihood ratio tests indicate that the best-fitting model for BS speakers comprises linear and quadratic terms, while for LP speakers, it includes only the linear term. Conversely, for long vowels, the optimal model for BS speakers incorporates only the linear term, whereas for LP speakers, it includes both linear and quadratic terms.

The findings reveal that Tone 1 consistently exhibits a higher mean f0 than Tone 2 across both short and long vowels for speakers from both villages. The effect size ranged approximately from 0.1 to 0.6 z-scores. Additionally, interactions were observed between the linear term and tone, indicating distinct slopes between the two tones, for both vowel length and for speakers from both villages. Specifically, Tone 2 demonstrates a steeper rising slope. Furthermore, interactions between the quadratic term and tone were observed for short vowels of BS speakers and long vowels of LP speakers, suggesting differences in the curvature of the f0 trajectories. Figure 1 illustrates the f0 trajectories predicted.

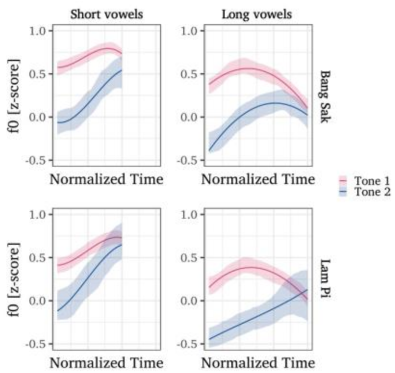

Figure 1: *GCA predicted f0 trajectories over the vowels of short (left) and long vowels (right) spoken by speakers from BS (top) and LP (bottom). Ribbons represent Standard Errors.*

[1] Other spectral measures, such as H1*-H2*, were also extracted, but they did not attain statistical significance in the models.

3.2. F1

In the case of short vowels, the likelihood ratio tests indicate that the best-fitting model for BS speakers comprises of only the intercept term for speakers from both villages. Conversely, for long vowels, the optimal model for BS speakers incorporates linear, quadratic, and cubic terms, whereas for LP speakers, it includes linear and quadratic terms.

The findings reveal that Tone 1 consistently exhibits a higher mean F1 than Tone 2 across both short and long vowels for speakers from both villages. The effect size ranged approximately from 0.02 to 0.5 z-scores. Additionally, interactions were observed between the linear term and tone, indicating distinct slopes for the two tones, for long vowels of LP speakers. Furthermore, interactions between the quadratic term and tone were observed for long vowels of BS speakers and of LP speakers, suggesting differences in the curvature of the F1 trajectories across tonal categories. The significance of interaction between cubic term and tone for long vowels of BS speakers also indicate differences in the curvature of the F1 trajectories.

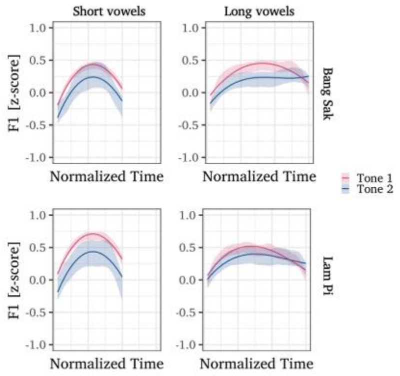

Figure 2: *GCA predicted F1 trajectories over the vowels.*

3.3. F2

In the case of short vowels, the likelihood ratio tests indicate that the best-fitting model for BS speakers comprises linear, quadratic, and cubic terms, while for LP speakers, it includes only the intercept term. Conversely, for long vowels, the optimal model for BS speakers incorporates linear, quadratic, and cubic terms, whereas for LP speakers, it includes linear and quadratic terms.

The findings reveal that Tone 1 consistently exhibits a slightly higher mean F2 than Tone 2 across both short and long vowels for speakers from both villages. The effect size ranged approximately from 0.02 to 0.09 z-scores. Additionally, interactions were observed between the linear term and tone, indicating distinct slopes between the two tones, for long vowels of speakers from both villages. Furthermore, interactions between the quadratic term and tone were observed for long vowels of speakers from both villages, suggesting differences in the curvature of the F2 trajectories across tonal categories. The significance of interaction between cubic term and tone for long and short vowels of BS speakers also indicate differences in the curvature of the F2 trajectories.

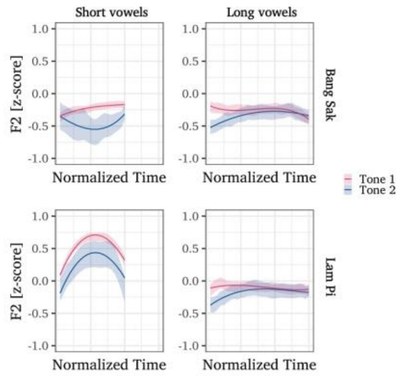

Figure 3: *GCA predicted F2 trajectories over the vowels.*

3.4. H1*-A3*

In the case of short vowels, the likelihood ratio tests indicate that the best-fitting model for speakers from both villages comprises only the linear term. Similarly, for long vowels, the optimal model for speakers from both villages incorporates only the linear term as well.

The findings reveal that Tone 1 consistently exhibits a lower mean H1*-A3* than Tone 2 across both short and long vowels for speakers from both villages. The effect size ranged approximately from -0.04 to -0.2 z-scores. Additionally, interactions were observed between the linear term and tone, indicating distinct slopes between the two tones, for long vowels of speakers from both villages.

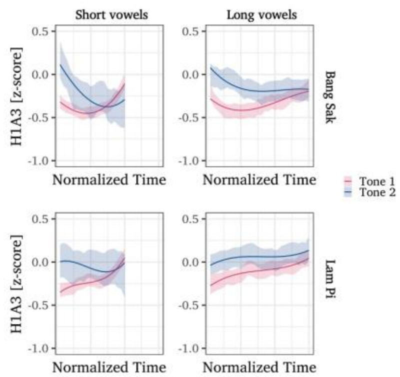

Figure 4: *GCA predicted H1*-A3* trajectories over the vowels.*

3.5. CPP

In the case of short vowels, the likelihood ratio tests indicate that the best-fitting model for BS speakers comprises linear and quadratic terms, while for LP speakers, it includes linear and quadratic terms. Conversely, for long vowels, the optimal model for speakers from both villages incorporates linear and quadratic terms.

The findings reveal that Tone 1 consistently exhibits a slightly higher mean CPP than Tone 2 across both short and long vowels for speakers from both villages. The effect size ranged approximately from 0.2 to 0.9 z-scores. Additionally, interactions were observed between the linear term and tone, indicating distinct slopes between the two tones, for long vowels of BS speakers and short vowels of LP. Furthermore, interactions between the quadratic term and tone were observed for short and long vowels of speakers from both villages, suggesting differences in the curvature of the CPP trajectories across tonal categories.

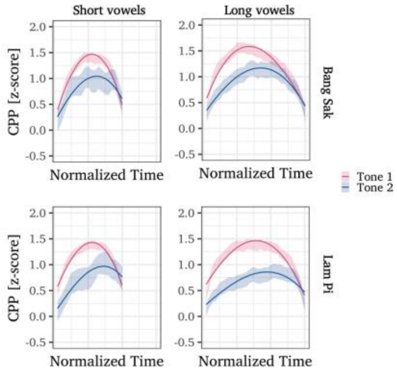

Figure 5: *GCA predicted CPP trajectories over the vowels.*

Furthermore, we observed significant effects of onset voicing on various acoustic measures, including the mean and slope of f0, the mean, slope, and curvature of F1, the mean and slope of F2, the mean and slope of H1*-A3*, and the mean, slope, and curvature of CPP. Additionally, coda class displayed significant effects on the mean and slope of f0, the mean and slope of F1, the mean, slope, and curvature of F2, the mean and slope of H1*-A3*, as well as the mean, slope, and curvature of CPP. Lastly, we also noted interactions between onset voicing and tone across all measures.

4. Discussion

Our experimental findings strongly suggest that f0 serves as the primary acoustic cue for tonal contrast in Moklen, accompanied by differences in vowel quality and phonation type. Specifically, Tone 1 is characterized by a higher pitch and a lower, more front vowel with modal voice, while Tone 2 exhibits a lower pitch and a higher, more back and, breathier vowel. These characteristics bear similarities to register distinctions observed in Austroasiatic and Chamic languages of Mainland Southeast Asia [18], suggesting a potential transphonologization of laryngeal properties into prosodic ones in Moklen. Register distinctions in these languages involve a bundle of acoustic properties (f0, phonation type, and vowel quality) that are realized on the rhymes, but they originated from a voicing contrast in onsets [19].

Despite these similarities, Moklen tones do not appear to emerge from the same process as in Austroasiatic languages. While most Austroasiatic languages developed tones through the loss of onset voicing contrast, Moklen still retains the onset voicing of the proto language. Crucially, tones in Moklen cannot be predicted solely from onset voicing, as evidenced by words like /ʔáːk/ ‘top,’ /ʔàːk/ ‘crow,’ /kabáːŋ/ ‘boat,’ /padàː/ ‘spur,’ /dabàːk/ ‘short,’ etc., which encompass all possible combinations of onset voicing and tones. Thus, the possibility that Moklen tones originated from laryngeal properties of onset voicing is ruled out.

The origin of Moklen tone remains a perplexing mystery, lacking empirical evidence or in-depth study, despite proposed hypotheses in previous research. These hypotheses include the influence of Southern Thai, an adjacent tonal language [1], [4], [5], [6], and additional segmental sources such as preceding syllables [6]. Besides the proposals of contact-induced tonogenesis, it is uncertain whether this development occurred as a result of external contact or an internal process, which could have involved additional segmental sources, such as preceding syllables, or other prosodic elements, such as stress positions, as observed in Southern Qiang dialects [20].

As a final note, it is worth pointing out that while pitch, vowel quality, and phonation type serve as phonetic cues in Moklen tonal distinction, it remains unclear whether these cues are used by native Moklen listeners in perception. Further investigation is required to provide insights into this aspect.

5. Conclusions

This pioneering acoustic study on Moklen tones sheds light on the phonetic differentiation of these tonal categories. Our results, measured from f0, F1, F2, H1*-A3*, and CPP, demonstrate that f0 is not the sole acoustic cue for tonal contrast in Moklen; vowel quality and voice quality also play significant roles. We observed consistent differences between Tone 1 and Tone 2 across short and long vowels for speakers from both BS and LP villages. Specifically, Tone 1 has higher F0, while Tone 2 exhibits steeper F0 slopes. Both tones also display differences in the curvature of the f0 trajectories. Regarding vowel quality, Tone 1 has a higher mean F1 and slightly higher mean F2 compared to Tone 2. In terms of voice quality, Tone 1 exhibits a lower mean H1*-A3* and a slightly higher mean CPP compared to Tone 2 across both short and long vowels. These findings underscore the distinct phonetic characteristics between Tone 1 and Tone 2 in Moklen. In conclusion, Tone 1 is characterized by higher pitch and a lower, more front vowel with modal voice, while Tone 2 has a lower pitch and a higher, more back, and breathier vowel. They significantly contribute to our understanding of tonal contrasts in Moklen and lay the groundwork for further research into its tonal evolution.

6. Acknowledgements

This research was part of the project “Research and Documentation of the Moklen Language and Culture in the Southeast Asian Context” supported by the Institute of Suvarnabhumi Studies, Thailand Academy of Sciences, Humanities, and Arts, and the Ministry of Higher Education, Science, Research, and Innovation. Additionally, heartfelt thanks are extended to the TAI 2023 Organizing Committee, Chulalongkorn University, Ramkhamhaeng University, and the Institute for Phonetics and Speech Processing (IPS) for their invaluable support. Furthermore, we would like to express our appreciation to the three anonymous reviewers for providing their constructive and useful comments. Special thanks are also extended to all Moklen participants for their invaluable contributions to the research project.

7. References

[1] M. D. Larish, “Moken and Moklen,” in *The Austronesian Languages of Asia and Madagascar*, K. A. Adelaar and N. Himmelmann, Eds., Routledge, 2005.

[2] S. Premsrirat, “Language situation in Thai society and ethnic diversity,” *Journal of Language and Culture*, vol. 25, no. 2, pp. 5–17, 2006.

[3] P. Pittayaporn, W. Pornpottanamas, and D. Loss, Eds., *Moklen-Thai-English dictionary: a pilot version*.

Bangkok: Academic Work Dissemination Project, Faculty of Arts, Chulalongkorn University, 2022.
[4] P. Swastham, "A description of Moklen: A Malayo-Polynesian language in Thailand," Master's thesis, Mahidol University, 1982.
[5] M. D. Larish, "Moklen-Moken Phonology: Mainland or Insular Southeast Asian Typology?," in *Proceedings of the Seventh International Conference on Austronesian Linguistics*, Leiden: Rodopi, 1997, pp. 125–149.
[6] M. D. Larish, "The position of Moken and Moklen within the Austronesian language family," Ph.D. dissertation, University of Hawai'i at Manoa, 1999.
[7] A.-G. Haudricourt, "De l'origine des tons en vietnamien," *Journal Asiatique*, vol. 242, pp. 69–82, 1954.
[8] P. Pittayaporn, "On becoming Mainland: Unraveling Malay influence on Moklenic languages," *Sojourn*, vol. 39, no. 1, in press.
[9] J.-O. Svantesson and D. House, "Tone production, tone perception and Kammu tonogenesis," *Phonology*, vol. 23, no. 02, pp. 309–333, Aug. 2006, doi: 10.1017/S0952675706000923.
[10] M. Brunelle, "A phonetic study of Eastern Cham register," in *Chamic and beyond: Studies in mainland Austronesian languages*, Pacific Linguistics, 2005.
[11] J. Edmondson and N. V. Lợi, "Tones and voice quality in modern northern Vietnamese: Instrumental case studies," *Mon-Khmer Studies*, vol. 28, pp. 1–18, 1998.
[12] J. W. Watkins, "Burmese," *Journal of the International Phonetic Association*, vol. 31, no. 2, pp. 291–295, 2001, doi: 10.1017/S0025100301002122.
[13] J. F. Gruber, "An articulatory, acoustic, and auditory study of Burmese tone," Ph.D. dissertation, Georgetown University, 2011.
[14] M. Brunelle and J. Kirby, "Tone and Phonation in Southeast Asian Languages," *Language and Linguist. Compass*, vol. 10, no. 4, pp. 191–207, Apr. 2016, doi: 10.1111/lnc3.12182.
[15] J. P. Kirby, "PraatSauce: Praat-based tools for spectral analysis." 2021. [Online]. Available: https://github.com/kirbyj/praatsauce
[16] D. Hirst, "The analysis by synthesis of speech melody: from data to models," *Journal of Speech Sciences*, vol. 1, no. 1, pp. 55–83, 2011, doi: 10.20396/joss.v1i1.15011.
[17] D. Mirman, *Growth Curve Analysis and Visualization Using R*. Boca Raton: Chapman & Hall/CRC the R series, 2014.
[18] M. Brunelle and J. Kirby, "Re-assessing tonal diversity and geographical convergence in Mainland Southeast Asia," in *Languages of Mainland Southeast Asia: The state of the art*, De Gruyter Mouton, 2015, pp. 82–111.
[19] M. Brunelle, T. T. Tấn, J. Kirby, and Đ. L. Giang, "Transphonologization of voicing in Chru: Studies in production and perception," *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, vol. 11, no. 1, pp. 1–33, Oct. 2020, doi: 10.5334/labphon.278.
[20] J. P. Evans, "Contact-induced tonogenesis in Southern Qiang," *Language and Linguistics*, vol. 2, no. 2, pp. 63–110, 2001.