

**Reanalyzing Thai Glottal Codas:
Probability, perception, and displaced contrast**

J. Fraser Bennett

**SIL International
2005**

Contents

Abstract

1. Introduction
2. The phonological status of ?
 - 2.1 Previous analyses
 - 2.1.1 Prosodic, structural, and early generative analyses
 - 2.1.2 Bennett's (1995, 1998) metrical analyses
 - 2.2 On uniform underlying segment inventories and the underlying status of ?
3. Predictable glottal codas as displaced vowel length markers
 - 3.1 Preliminaries
 - 3.2 Methods
 - 3.2.1 Estimating probabilities of miscategorization
 - 3.3 Predictions and data
 - 3.4 Establishing the contextual frame for categorization
4. Tabulated data
5. Calculated probabilities of miscategorization
6. Conclusions

Appendix 1

Postscript on formalizing the analysis

Appendix 2

Kolmogorov-Smirnov goodness-of-fit test results (.05 confidence level)

Appendix 3

Outline of Bennett's (1995) optimality-theoretic analysis of stress placement and glottal epenthesis

References

Abstract

The distribution of glottal stop codas in Thai is largely predictable, suggesting that they ought to be absent from lexical representations. However, tonal patterns and a handful of exceptional forms seem to require that they be present. This paper proposes a novel approach to this well-known problem, namely, that glottal codas are neither fully absent nor fully present: word-final glottal codas may be underlyingly present, but those in other environments are inserted. A motivation for this epenthesis is suggested by an analysis of the phonetic duration and textual frequency of various syllable types, which indicates that the perceptual reliability of the vowel length contrast would likely be seriously compromised in stressed, open syllables. The glottal coda of a CV? syllable may thus act as a displaced contrast marker for the preceding vowel's shortness.*

1. Introduction

Glottal stops in coda position are, with very few exceptions, not contrastive in Standard Thai, being limited to stressed syllables with short vowels. Although a handful of words (mostly loanwords) contain a falling diphthong followed by a glottal coda, glottal codas are never found after long vowels or rising diphthongs.

Despite this predictable distribution, the underlying status of glottal codas has long been problematic. Their predictable distribution would usually indicate that glottal codas should be absent from underlying representations. However, the fact that syllables with obstruent codas—oral or glottal—show identical restrictions in the range of tones they may bear has led most researchers to treat glottal codas as present in the lexical specifications of morphemes (Henderson 1949, Haas 1964, Udom 1967, Aporn 1973, Duryea 1991).

This paper attempts to clarify the status of glottal codas by bringing two sorts of data to bear on the problem. First, additional phonological evidence is presented to argue that some, but not all, glottal codas are present underlyingly, and that the distribution of glottal codas in underlying representations closely matches their distribution in “normal” or casual speech styles. Second, phonetic and textual evidence is presented to suggest a contrastive motivation for the distribution of glottal codas. A preliminary analysis of the probability that listeners will miscategorize vowel lengths in various syllable types showed that, on the basis of phonetic duration and textual frequency, a stressed, open CV syllable is very likely to be misinterpreted as having a long vowel (i.e., CV:). Few open, stressed [CV:] syllables are likely to be misinterpreted as [CV], however, and the same disparity is not found with other syllable types. The distribution and behavior of coda glottal stops in Standard Thai may therefore best be viewed as a displaced contrast, with the glottal codas enhancing an otherwise weak contrast between long and short vowels in open syllables

*The author would like to thank Paul Boersma, Mike Cahill, Paulette Hopple, and Keith Slater, and participants at the “Role of Perceptual Phenomena in Phonology” satellite conference of the 14th International Congress of Phonetic Sciences (San Francisco, 1999) for comments and discussion on earlier versions of this paper; also the Graduate Institute of Applied Linguistics, Dallas, TX, for leave time in support of this research, and Peter Martin for unparalleled technical help. All errors of fact and logic are naturally my own.

Section 2 briefly describes the distribution of glottal stops in Standard Thai, reviews relevant previous analyses (with particular attention to Bennett’s (1995, 1998) metrical analysis of glottal epenthesis), and presents the phonological evidence for a “what-you-see-is-what-you-get” distribution of underlying glottal codas. Section 3 takes up the question of what might motivate the epenthesis of glottal codas in non-final syllables, outlines the displaced contrast hypothesis and presents phonostatic and phonetic data that short vowels in stressed, open syllables are highly susceptible to being miscategorized. Section 4 discusses the import of these results for the phonological analysis of the Thai metrical system, and proposes a framework for formalizing the epenthesis.

I will deliberately avoid casting the argument in any particular theoretical framework, beyond the minimum assumptions that underlying representations exist and that surface-level restrictions can crucially affect the mapping between underlying and surface forms. However, I will propose the rough outlines of a formalization for the epenthetic mechanism using an optimality-theoretic vocabulary (§4.2).

2. The phonological status of ?

2.1 Previous analyses

A thorough account of the principles which determine the consonant and vowel inventories and possible syllable shapes is well beyond the range of this study, but a brief review of the Thai consonant and vowel systems should set the stage for our consideration of glottal codas. The inventory of Thai surface segments appears in (1–2).

- (1) Consonants
 Onsets: $p^h, t^h, t\phi^h, k^h, p, t, t\phi, k, \text{ʔ}, b, d, m, n, \eta, f, s, h, w, j, l, r$
 Codas: $p, t, k, \text{ʔ}, m, n, \eta, w, j$
- (2) Vowels
 Short: $i, u, \text{u}, e, \text{ɤ}, o, \text{æ}, a, \text{ɔ}$
 Long: $i:, u:, \text{u}:, e:, \text{ɤ}:, o:, \text{æ}:, a:, \text{ɔ}:$
 Falling diphthongs: $i\text{ə}, u\text{ə}, u\text{ə}$

Stress-bearing syllables adhere to the template in (3).

$$(3) \quad \begin{array}{c} \text{T} \\ C_i(C_m)V_1(V_2)(C_f) \end{array}$$

The nucleus V_1 (V_2) may be composed of a short vowel, a long vowel, or a diphthong. There may be single coda consonant.² However, the rime in a stressed syllable may not be a single short vowel. It must

¹ Glottal stops in onset position are unstable, being present in carefully articulated speech but readily elided in more rapid, casual styles. Thus, one finds both [ʔa.raj] and [a.raj] ‘what?’, [ʔi:k] and [i:k] ‘another’, etc. This alternation appears to be of the sort which is straightforwardly handled in a procedural model by variable rules, and I will not be much concerned with it.

² A number of further restrictions apply: for instance, V_1 and C_f may not have the same place of articulation.

be at least a long vowel, or a short vowel plus a coda.³ The minimal stressed syllable is therefore not CV, but CVX, where X represents a consonant, length in a long vowel, or the second half of a diphthong.

Syllables that bear secondary stress are reduced in duration and tonal contours (Siripong, Gandour and Harper 1994, 1998), but otherwise follow the pattern of (3). Fully stressless syllables are much more restricted in form, being limited to the pattern in (4), where V is one of the three cardinal vowels, [i], [a], or [u], but generally centralized toward [ə].

(4) C_i (C_m) V

- (5) a. *prə.kà:t* ‘notice, announcement’
 b. *hi.ˈmáʔ* ‘snow’
 c. *bu.ˈrùt* ‘sp. monster’

In coda position, glottal stops are generally found only after short vowels in stressed syllables. Glottal codas never appear on unstressed syllables.

- (6) a. *dùʔ* ‘fierce’
pàʔ ‘to patch’
 b. *hi.ˈmáʔ* ‘snow’
tʰə.mə.ˈnáʔ ‘ascetic’
 c. *pàʔ.ja:ŋ* ‘to patch a tire’

A handful of words with glottal stops following falling diphthongs are found: Haas (1964) lists three (7a–c), to which a fourth can be added.

- (7) a. *pʰùəʔ.pʰiəʔ* ‘imitative of the sound of beating, slapping, cracking’
pʰúəʔ.pʰiəʔ
 b. *ˈta:m ˈdiəʔ* ‘to follow closely’
 c. *ˈtɕiəʔ* ‘to eat’ (slang)
 d. *pɔʔ.píəʔ* ‘spring roll’

Such words are quite rare, however. Glottal codas appear not to occur at all after long vowels or rising diphthongs.⁴

In addition to these facts, any analysis of glottal codas must account for the variation between “normal” and hyperarticulated speech styles—what Henderson (1949) called “combinative” and “isolative” styles,

³ There is some uncertainty over the proper treatment of surface triphthongs such as [tɕiəŋ] ‘deep fat fried’, [dùəj] ‘also’: should the final glide be treated as a final consonant or as a vocalic element? The former case simplifies the syllable template, the latter the set of possible codas. As the issue is not crucial for the analysis of glottal codas, I will not pursue it.

⁴ Rungpat Roengpitya (p.c.) points out that the feminine politeness particle *kʰâ:* is often rendered in casual speech as *hâ:ʔ*. The phonology of sentence particles diverges in a number of ways from the rest of the phonology, and I will not try to account for them here.

respectively.⁵ “Normal” speech styles allow light, stressless CV syllables to appear in non-final position. But in the hyperarticulated style, these syllables bear primary stress and appear with a glottal coda.

(8)	“Normal”	Hyperarticulated	(‘Isolative’)
	<i>ma.ˈna:rw</i>	<i>máʔ.ˈna:rw</i>	‘citron’
	<i>ˈtʰə.ˈbiən</i>	<i>ˈtʰáʔ.ˈbiən</i>	‘to register’
	<i>kə.rə.kə.ˈda:</i>	<i>kàʔ.ˈráʔ.kàʔ.ˈda:</i>	‘July, (zodiacal) Cancer’

2.1.1 Prosodic, structural, and early generative analyses

Despite the predicable distribution of the glottal codas, most researchers have treated them as an essential part of phonological specifications, rather than as epenthetic elements (Henderson 1949, Haas 1964, Udom 1967, Aporn 1973).⁶ Aporn carefully justifies this treatment on the basis of the tones that may appear on CVʔ syllables. Like other obstruent-final syllables with short vowels (hereafter CVS, where S represents a surd), CVʔ syllables may carry only high or low tone, never mid, falling, or rising tone. Aporn takes this as evidence for a morpheme structure condition that defines the range of possible tones on CVS syllables. Since the morpheme structure conditions of early generative phonology apply to underlying representations only, Aporn concludes that glottal codas must be present in underlying representations. These glottals surface in the hyperarticulated style, but in other speech styles they are removed by a rule that deletes glottal codas in unstressed syllables. (9) reproduces Udom’s (1967:564) glottal deletion rule.

$$(9) \quad ? \rightarrow \emptyset / \left\{ \begin{array}{l} _[-prom] \\ [-prom]_ \\ [+prom]_ \$ \end{array} \right\}$$

A glottal stop at syllable final is deleted if it follows or precedes a nonprominent [i.e., stressed—JFB] vowel; or follows a prominent vowel and precedes a syllable boundary.

These analyses and others like them (Apiluck 1993, Duryea 1991) account for the range of tones observed on CVʔ syllables, and for the $\emptyset \sim ?$ alternation between normal and hyperarticulated speech styles in non-final syllables. However, they do not explain the extreme rarity of coda glottals after falling diphthongs, nor their total absence after long vowels. Moreover, they posit underlying representations littered with glottal stops that surface only in a relatively rare hyperarticulated speech style and that have no appreciable effect on phonetic forms in normal speech styles—no easy conclusion for a learner to draw.

2.1.2 Bennett’s (1995, 1998) metrical analyses

Bennett (1995, 1998) sought to remedy these difficulties by proposing that glottal codas are not present at all in underlying representations, but are epenthetic. Bennett (1995) claimed that glottal codas are inserted in order to satisfy two sorts of metrical requirements: first, a set of requirements that place primary word

⁵ Henderson (1949) distinguished three styles: *isolative*, *combinative*, and *rapid combinative*, the last being a casual, hypoarticulated style. The distinction between combinative and rapid combinative styles will not concern us here, and the examples given in the text are from the middle, combinative style.

⁶ Henderson’s Firthain analysis treated glottal codas as a conjunction of two prosodic features, *plosion* and ζ ‘zero’ (in this context, no oral consonant place). The plosion feature is omitted in close juncture environments, leaving the zero. In Haas’ structuralist analysis, glottal stops are bona fide phonemes, realized by a zero allophone under conditions of stresslessness. The early generative analyses of Udom (1967) and Aporn (1973) treat glottal codas as underlying phonemes that are deleted in unstressed environments.

stress on the word-final syllable; and second, the *Stress-to-Weight Principle* (SWP), which requires all stressed syllables to be metrically heavy. When an underlyingly light syllable occurs in word-final position, it receives primary stress; in order to satisfy the SWP, it is augmented with a glottal coda (10a). The differences between the normal and hyperarticulated styles are traced to metrical differences between the two speech styles. In the hyperarticulated ‘isolative’ style, each syllable is treated as a distinct prosodic word and receives its own primary stress, with the result that underlying /CV/ syllables are augmented with ʔ (10b). In normal speech styles, however, these CV syllables are allowed to surface without stress.

Bennett (1995) is cast in terms of Optimality Theory, with constraints rather than phonological rules mapping between underlying and surface forms. This lets surface-level constraints restrict the range of tones on CVʔ syllables, so that ʔ need not be present in underlying representations for the sake of morpheme structure constraints. (See appendix 3 for details of Bennett’s (1995) optimality theoretic formulation.) In an equivalent rule-based analysis, the following rules would apply.

(10)	a.	Normal style	b.	Isolative style	
	Underlying representation	/t ^h úrá/		/t ^h úrá/	‘affairs’
	Stress assignment	t ^h ú . 'rá		t ^h ú . 'rá	
	Glottal epenthesis	t ^h ú . 'ráʔ		t ^h úʔ . 'ráʔ	
	Tonal reduction	[t ^h u . 'ráʔ]		[t ^h úʔ . 'ráʔ]	

In addition, Bennett (1995) proposes to connect the lack of final light syllables with the overwhelming number of iamb-shaped words in the lexicon. (Iambs are defined according to Hayes (1995), as either (—) or (• —), where • represents a short syllable and — represents a long syllable.) As for the SWP, which forces glottal epenthesis, Bennett claims that it is one of three constraints that define iambic foot structure itself.⁷ Thus, rather than showing “monosyllabic” tendencies, as commonly claimed, the Thai lexicon tends toward *monopody*—that is, words that are one foot in size. On this analysis, the principle that forces the epenthesis of glottal codas in stressed syllables comes “for free” from the iambic footing system.

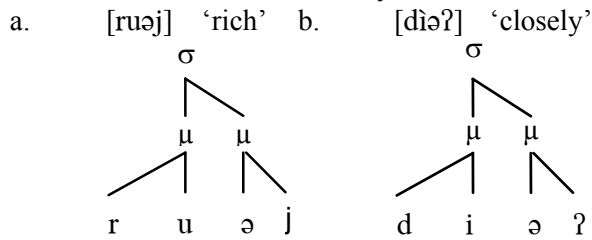
There are, however, serious weaknesses to Bennett’s metrical proposal. First, it has no explanation for the handful of attested CV₁V₂ʔ syllables. There is reason to think that CV₁V₂ʔ syllables such as [d̪iəʔ] are bimoraic, with the same moraic structure as other CV₁V₂C syllables such as [r̪uəj] ‘rich’ (11).⁸ But if *all* glottal codas are epenthetic, then structures like (10b) ought to be underivable.

⁷ The other constraints that define iambic structure in Bennett (1995) are *Rhythm Type-Iambic* and *Foot Binarity*.

- i. Rhythm Type-Iambic Align (Ft, R, Head(Ft),R)
The right edge of every foot is aligned with the right edge of the head of a foot—i.e., feet are stressed at their right edge.
- ii. Foot Binarity Feet are binary in terms of μ or σ.
(Note that monosyllabic (—) iambs satisfy the constraint by virtue of being bimoraic, (• —) iambs by virtue of being disyllabic.)

Other proposals for constraints to define iambic footing include Kager (1993), Hung (1993), Suzuki (1995), and Eisner (1997). I will not pursue any comparison between them here, but further developments—in particular, Bennett (1998) and the research reported here—suggest that Bennett’s 1995 proposal may not be tenable.

⁸ Broselow, Chen, and Huffman (1997) report statistically significant correlations between vowel interval length and moraic (Footnote continued on next page)

(11) Metrical structure of CV₁V₂C syllables

Second, the Stress-to-Weight Principle, which forces the epenthesis of glottal codas on underlyingly light syllables in stressed positions, is suspect. Bennett (1995) appeals to iambic foot structure as independent evidence for the SWP. However, the placement of secondary stresses in Thai depends not on foot structure, but on metrical prominence. Non-final heavy syllables bear secondary stress, while non-final light syllables are stressless (12).⁹ This is in stark contrast to the usual pattern of foot-driven stress systems, which typically assign stress to every other mora or syllable.¹⁰

(12)	$\bar{\cdot}$. $\bar{\cdot}$	<i>ban</i> . $\bar{\cdot}$ <i>daj</i>	'ladder'
	$\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$	<i>na:</i> . <i>li</i> . $\bar{\cdot}$ <i>ka:</i>	'timepiece'
	$\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$	<i>nə</i> . <i>li</i> . $\bar{\cdot}$ <i>ni:</i>	'lotus pond'
	$\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$	$\bar{\cdot}$ <i>sǎ:</i> . $\bar{\cdot}$ <i>t^ha:</i> . <i>rə</i> . $\bar{\cdot}$ <i>náʔ</i>	'common'
	$\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$	<i>kə</i> . <i>rə</i> . <i>kə</i> . $\bar{\cdot}$ <i>da:</i>	'July/(zodiacal) Cancer'
	$\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$. $\bar{\cdot}$	<i>p^hít</i> . <i>sə</i> . <i>nu</i> . $\bar{\cdot}$ <i>lô:k</i>	'Phitsanulok' (place name)

structure in Hindi, Malayalam, and Levantine Arabic, according to the moraic structures of each language. The vowel intervals reported by Siripong et al. (1998) for Thai are similar to those of Arabic, in which CV: and CV:C are bimoraic.

⁹ On prominence-driven stress systems, see Prince and Smolensky 1993, Hayes 1994, Walker 1997.

¹⁰ An analysis which incorporates metrical feet encounters further problems in the parsing of words that contain strings of two or more light syllables. It is impossible to tell whether such strings of light syllable should occupy weak positions in an unbounded foot (i) or remain unparsed (ii).

i.	[(kə . rə . kə . $\bar{\cdot}$ da:)] 'July'	[(p ^h ít)(sə . nu . $\bar{\cdot}$ lô:k)]	'Phitsanulok'
ii.	[kə . rə . (kə . $\bar{\cdot}$ da:)]	[(p ^h ít) sə . (nu . $\bar{\cdot}$ lô:k)]	

In constraint-theoretic terms, there is an indeterminate ranking between two conflicting constraints, *Parse- σ* (which favors exhaustive footing) and *Foot Binariness* (which favors strictly binary feet). This situation is illustrated in constraint tableau (iii), where the ranking of Foot-Binariness and Parse- σ is left indeterminate.

iii. Indeterminate ranking between Parse- σ and Foot Binariness

	/p ^h ít.sə.nu.lô:k/ (Normal style)	End Rule-R WSP	Parse- σ	Ft-Bin
? (i)	[(p ^h ít)(sə.nu.lô:k)]			*
? (ii)	[(p ^h ít) sə (nu.lô:k)]		*	

Although each constraint favors a different prosodic structure, the nature of the data makes it impossible to determine the ranking. Whether this is a situation that Optimality Theory can tolerate is not clear, but it is a troublesome feature for the formal analysis. (For further details on the analytic indeterminacies introduced by the metrical analysis, see Bennett 1998). In fact, it appears to make no difference which parsing we choose.

In fact, ignoring for the moment the word-final glottal stops of (6), an adequate analysis of Thai stress could be framed in terms of two prominent constraints: the *Weight-to-Stress Principle* (WSP; Prince 1990), which requires heavy syllables to be stressed; and *End Rule-Right* (Hayes 1994), which places main word stress on the last syllable of the word.

Without independent evidence for the SWP, its role in the grammar is tenuous. The SWP is not widely attested; Prince (1990), for instance, discounts it as an “active principle of grammar.” While we might discard any notion of iambic foot structure and still retain the SWP as a motivation for glottal epenthesis, the wider account of Thai prosody would lack the explanatory connections between stress placement, glottal epenthesis and lexical monophony. The metrical analysis is thus torn between simplicity and scope of explanation.

Finally, although Bennett succeeds in purging predictable glottal codas from underlying representations, the shift in explanatory focus from representations to surface-level constraints, characteristic of Optimality Theory, makes this a result of dubious value. In sum, the difficulties associated with the analyses proposed thus far are severe enough to warrant a reconsideration of glottal codas and their role in Thai phonology.

2.2 On uniform underlying segment inventories and the underlying status of ?

A major stumbling block to previous analyses has been the assumption that glottal codas must be present in all environments underlyingly, or absent in all environments. Such uniformity is not necessary, however, if we adopt an analytic framework that includes strong surface-level constraints. The principal reason for underlying glottal codas in Udom (1967) and Aporn (1973) was to ensure that CV? syllables have the same range of tones as CVS syllables. But surface constraints can do this work, without reference to underlying forms.

We are therefore free to consider each environment in which glottal codas appear in its own right. In the remarks that follow, I will also avoid the assumption that the SWP plays any crucial role in the distribution of ?, in the absence of independent evidence that the SWP plays a role in the grammar.

First, we may observe that glottal codas must be in the lexical specification (= underlying representation) of CV₁V₂? words, since these are truly unpredictable.¹¹

Second, we may note that epenthesis of glottal codas in stressed syllables is independently required for loanword phonology, in order to account for the numerous borrowings from Sanskrit and Pali that have acquired a glottal coda as part of their adoption into Thai.¹²

- (13) *hi.ˈmá?* ‘snow’ < *hima*
 kʰa.ˈtì? ‘rebirth’ < *gati* ‘pass by, elapse’
 tʰu.ˈrá? ‘affairs’ < *dhura*

The ? ~ Ø alternation between isolative and normal speech styles has essentially the same environment as the epenthetic glottal codas of loan forms, namely, stressed CV(?) syllables. It would therefore be an analytic loss, no matter what framework we use, to account for them with two separate processes. Rather, the alternating glottal codas should be regarded as strictly epenthetic.

But precisely because these glottal codas do not alternate, the same strong surface constraints require us to be agnostic as to their underlying status. The surface constraints are powerful enough to force the epenthesis of the glottal, but they say nothing about whether the underlying form ought to have the glottal or not; as McCarthy and Prince (1995:33f.) put it, other, more general principles must decide the matter.

Before rushing to a formal analysis of the alternation, it is fitting to ask *why* glottal codas have the distribution they do. After all, this is not part of the historical heritage of Thai. Optimality theory provides a vocabulary for posing rigorous formal theories that can (we hope) capture significant functional motivations. So before asking *how* the language implements the distribution of glottal codas, it’s good to ask *why*, lest we end up with a clever but myopic analysis. I’ll return later to some of the details necessary to a formal analysis. Although I will not, ultimately, provide a full analysis, I hope to establish certain functional motivations that will establish a firm footing for a formal analysis.

¹¹ Another class of unpredictable glottal codas is found in certain compounds based on Indic loan words. Some roots manifest an idiosyncratic ? when standing as the first element of a compound:

- | | | | | |
|-----|---------------------|-------------------------|---------------------------------|---|
| (i) | Column A | Column B | Column C | |
| | <i>tʰu.ˈrá?</i> | <i>tʰu.rə.ˈkìt</i> | <i>tʰu.ˈrá?.ˈkɑn</i> | ‘affairs’ / ‘business’ / ‘administrative’ |
| | <i>ˌsʰĩn.lə.pà?</i> | <i>ˌsʰĩn.lá.pa.ˈkɑm</i> | <i>ˌsʰĩn.lə.pà?.wít.tʰa.jɑː</i> | ‘art’ / ‘work of art’ / ‘arts and sciences’ |

In general, however, Indic roots that appear with a glottal coda lose the coda in compounds; for instance, *phít.sə.nù?* ‘Vishu’ ~ *pʰít.sə.nu* + *ˈlô:k* ‘world-pervading’ (place name). Such cases may be analyzed as cases of allomorphy—necessary also for roots that vary between consonant-final forms in isolation and vowel-final forms in compounds, such as *ˈkɑm* ‘deeds’ ~ *ˌkɑm.mə* + *ˈkɑn* ‘committee’.

¹² Many other borrowings from Sanskrit/Pali forms have lost their final short vowels, e.g. *kɑm* < Pali *kamma* ‘deeds’; whether the final syllable of any particular loanword is lost or augmented is a vexed topic that I will not address.

3. Predictable glottal codas as displaced vowel length markers

Asking *why* inevitably leads to matters of function and purpose, or to a conclusion of accident. Assuming that the distribution of ? is not a matter of blind chance, what role does this largely redundant segment play in the phonological economy of Thai? I believe we can make some initial observations.

First, I assume that speakers strive, among other things, to maintain distinctions between contrasting categories. If the differences between the phonetic exponents of two contrasting categories become so obscured that the contrast is no longer reliably carried, and if the contrast is important to speakers, then speakers will resort to other means, phonetic or lexical, to communicate the contrast.

In the case at hand, the distribution of predictable glottal codas in Thai supports the contrast between short and long vowels in an environment where that contrast would otherwise be unreliably perceived. In particular, the vowel length contrast is vulnerable in stress-bearing syllables that lack an oral coda. Presumably a glottal stop is suited for this task because the glottal articulation cleanly cuts off the vowel formants of the preceding vowel, and because it does not introduce any contrastive oral place features. Moreover, to the extent that the glottal coda is perceived as a distinct segment, it may become a displaced marker of vowel length in the minds of listeners. That is, the vowel length contrast is displaced from the vowel itself, to the presence or absence of a glottal coda. Let us call this the “displaced contrast hypothesis.”

The displaced contrast hypothesis makes two sorts of predictions. First, a perceptual study of listener reactions ought to show that listeners rely on the glottal coda as a marker of the preceding vowel’s length. Second, we would expect that phonetic measurements of the physical exponents of short and long vowels ought to show a significant overlap in stress-bearing syllables that lack an oral coda, to the point where short vowels would be regularly confused for long ones. Furthermore, since the displaced contrast marker, the glottal stop, appears after short vowels, we may tentatively assume that it is the short vowels that are more susceptible to misperception.

In this paper I will restrict myself to the second sort of evidence and leave a thorough perceptual study for future work. The results reported below should therefore be viewed as a preliminary study.

3.1 Preliminaries

Before proceeding, we must consider what the phonetic exponents of the long/short vowel distinction are. The present study focuses on duration differences between long and short vowels, but it is reasonable to ask whether vowel quality and phonation differences are sufficient to allow listeners to distinguish between long and short vowels, even if the differences in duration are reduced. Abramson (1960) reported significant differences in duration between “single” and “geminate” vowels in all nine vowel qualities; he also found small but measurable differences in vowel quality between long and short variants. In addition, it has been observed that syllables with obstruent codas are marked by noticeable glottalization of the vocalic nucleus.

The relative importance of vowel duration vis-à-vis vowel quality in conveying the vowel length contrast has not yet been established, but it appears that vowel quality alone cannot carry the distinction

(Abramson 1960:134ff., Rungpat Ruengpitya, p.c.).¹³ The functional load of glottalization in signaling vowel length is even less clear. We may note, however, that vowel glottalization is a feature of *all* obstruent-coda syllables, regardless of whether the coda is oral or glottal, and even of whether the vowel is long or short. This fact would indicate that the glottalization is a function of the obstruent coda, and not of vowel length per se. In the absence of further evidence, then, I will assume that vowel duration is the primary marker of the vowel length contrast.

For hearers to accurately categorize perceived vowels as long or short, they must at the very least perceive the start and end points of the vocalic articulation. Of course, phonetic stimuli are not identical, even when the stimuli encode the same category. Accurate perception of phonetic duration therefore does not guarantee accurate categorization: the listener must somehow resolve parametric durations into a categorial contrast. Following Boersma (1998), I assume that the phonetic instantiations of a single contrastive category—[a:], for example—are normally distributed around some mean vowel duration, quality, and phonation.¹⁴ The listener is thus faced with an exercise in probability: given a perceived vowel of a certain duration and spectral character, is the vowel more likely to belong to the population of short vowels or the population of long vowels? To this task the listener brings his or her implicit knowledge of the phonetic cues associated with each contrastive vowel length and the typical range of variation of each cue.

I will further assume that the listener also makes use of knowledge about the relative frequency with which each category occurs. Some categories are more likely to be encountered than others due to contrasting categories which occur in speech with different frequencies. This is due to their frequency in the lexicon and the rates at which various lexical items are used. Plainly all this is not conscious knowledge, but it seems reasonable to think that with long experience, the listener has an intuitive grasp of each of these parameters.

3.2 Methods

3.2.1 Estimating probabilities of miscategorization

Boersma (1998) presents a method for estimating the probability that any given percept will be categorized correctly or incorrectly. Given a pair of overlapping probability functions such as those in figure 1, their intersection (marked by the dashed vertical lines in figure 1) is first estimated. The area under each curve that lies beyond the critical point may then be calculated; the portion it represents of the entire area under the curve is the probability of miscategorization.

For instance, there should be various critical points along the spectrum of F_1 frequencies that listeners use to categorize perceived formants into vowel height categories. It is reasonable to think that these critical points are determined by the individual probabilities that any given F_1 frequency belongs to each vowel category—or more exactly, to each of the normally distributed populations of F_1 frequencies that instantiate each vowel height category. Other things being equal, the listener will classify any given

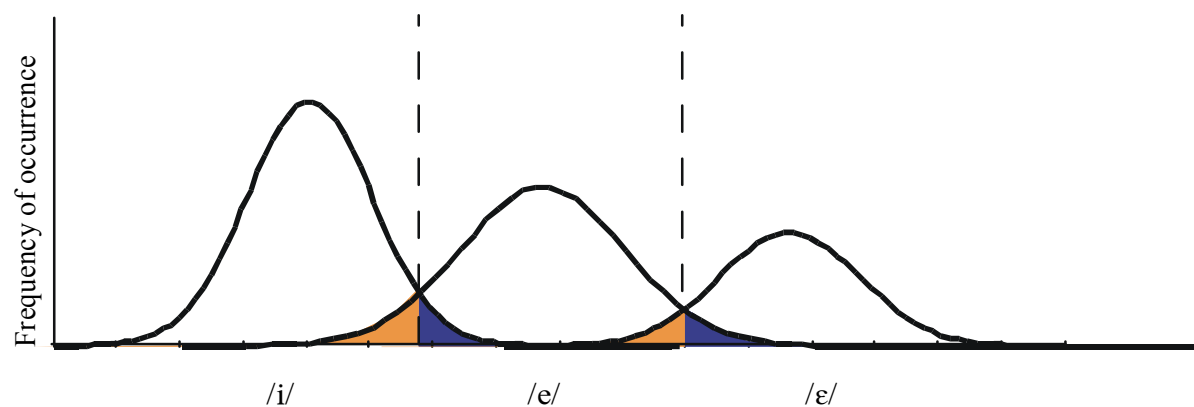
¹³ Abramson (1960:134ff.) found that subjects identified progressively shortened long vowel articulations as long down to, on average, 58 percent of the original long vowel's duration. Given that the durations of the short vowels in Abramson's experiment were, on average, 40 percent of the long vowel's duration, Abramson concludes that vowel quality alone cannot maintain the vowel length distinction. Compare Rungpat 1999, who found that the crossover point for shortened long syllables in nasal-final syllables was only 20 msec longer than a corresponding short vowel.

¹⁴ See appendix B for statistical tests for normality of vowel duration for one data set.

phonetic stimulus as an instance of the vowel height category that is most likely to be instantiated by the F_1 in question. In most cases this leads to a correct categorization, but instances along the tails of each population are likely to be misinterpreted as belonging to the next category over.

Figure 1 illustrates this scheme: each bell-shaped curve represents the distribution of the phonetic exponents of its vowel category. Their heights relative to each other reflect the relative frequency of their categories. Concretely, the normal distribution for each category is multiplied by a constant which represents its frequency, vis-à-vis other categories along the same phonetic dimension. The dashed vertical lines at the intersections of the probability curves indicate the critical values that listeners use in sorting F_1 percepts, the points at which there is a change in the most-likely category; the shaded portions of the tails indicate the range of F_1 values for which listeners are likely to miscategorize the F_1 percept.

Figure 1. Normally distributed F_1 populations instantiating three vowel height categories



In general, given a phonetic domain in which a particular category occurs more frequently than the others, all percepts within the domain will be assigned to that category. The probability that any particular percept will be assigned to that category is therefore the sum of the probabilities of a phonetic feature (here, vowel length) being produced within that domain, across all categories along the phonetic dimension.

$$(14) \quad P(f_{\text{perc}} = x_i | C_f = y) = \sum_i^n \left(\sum_a^b P(f_{\text{prod}} = x_i | C_f = y) \right)$$

where a and b are points along a phonetic dimension d such that for all points d_i such that $a \leq d_i \leq b$, $P(f_{\text{prod}} = x_i) \geq P(f_{\text{prod}} = x_j)$.

3.3 Predictions and data

If the displaced contrast hypothesis is true, this exercise in probability should yield unacceptably unreliable results for short vowels in stressed, codaless syllables. On the basis of vowel duration alone, there should be a high probability that listeners will incorrectly categorize short vowels in stressed, codaless syllables. In particular, the acoustic cues distinguishing long and short vowels, the range of typical variation in those cues, and the relative frequencies of occurrence of long and short vowels should combine to create a large probability that /CV/ and /CV:/ syllables will not be reliably distinguished. Given that the displaced contrast enhances short vowels rather than long vowels, the greater probability of miscategorization should lie in a particular direction: /CV/ ought to be more susceptible to being miscategorized as /CV:/ than vice versa. In addition, a high probability of confusion should not be found for other environments.

Using Boersma's method, the probabilities of misperception for short and long vowels in various environments were calculated, using two sets of data. The first data set is a composite of vowel duration data from Siripong, Gandour, and Harper (1998) and textual frequency data from McFarland (1942).¹⁵ Because there are significant correlations in Thai between vowel duration, vowel length, coda sonority, and stress (Siripong et al. 1998), separate calculations were made for each sonority type of coda and stress condition. The second set of data is a corpus of syllable duration measurements for a single text of 569 syllables reported in Theraphan (1977). The text, "Queen Srisawarinthira," contains a fairly high proportion of Indic loan words, where word-final glottals are well represented.¹⁶

3.4 Establishing the contextual frame for categorization

Categorization probability calculations should ideally meet two preconditions. First, vowel interval or syllable durations for each environment ought to be normally distributed. Siripong et al. (1998) present only summary results of their duration study, so we can only assume that their data are normally distributed. For the duration data reported in Theraphan (1977), appendix 2 summarizes the results of Kolmogorov-Smirnov goodness-of-fit tests for individual syllable type/stress conditions, as well as for aggregates of syllable type, stress condition and vowel length. For most coda types under primary stress, the Kolmogorov-Smirnov test results are consistent with normally distributed data.

Second, we must show that coda type and stress environment are actual conditioning environments for the physical duration of vowels. Tables 1 and 2 summarize the results of pairwise analyses of variance for various coda types and stress conditions.¹⁷ Under primary stress, coda types have a significant

¹⁵ Siripong et al. provide mean duration and standard deviation data for each of the syllable types (i.e., vowel length + coda sonority type) in the form of a figure, but do not provide the underlying figures. Numerical data used in the calculations here are estimated from their figure 2. Note also that because of the difficulty of segmenting running speech Siripong et al. do not distinguish CV? or CV: as separate categories. Although they suggest that codaless syllables pattern after obstruent-final syllables (which themselves are not distinguished from glide-final syllables in their data), tabulations of data from Theraphan (1977) (see below) cast doubt on this assumption. Nevertheless, I will follow this suggestion for the sake of using their otherwise excellent data.

McFarland (1942) provides a list of the 1,000 most common Thai words, based on a textual corpus of 167,546 words drawn from 30 distinct sources. (McFarland does not identify his sources, but certain of the words included indicate that at least some of the source texts are religious or mythological in character. Nor does McFarland describe the criteria used for making word breaks—not a trivial matter, since Thai orthography does not explicitly mark word breaks and compound words abound. However, the list seems, in my judgment, to be fairly representative.) Unfortunately, McFarland lists the words in alphabetical order, without any indication of their relative frequency; nevertheless, his list can serve as a first approximation, especially since it is based on a fairly large text corpus. Only word-final syllables were counted, in order to maintain consistency of stress environments.

¹⁶ Theraphan provides syllable duration measurements for all syllables in the text. While vowel interval duration data would be more to our purpose, two factors make it possible to use the syllable duration data. First, our major comparisons are within coda types, so that the inherent variation between the durations of different obstruent, glide, and nasal codas should be somewhat compensated. (Note, however, that Siripong et al. found a significant effect of syllable type on vowel duration.) On the other side of the syllable, the additional duration contributed by any medial consonants, which Theraphan (1977:69) finds to be around 50 msec, should be fairly randomly spread throughout the corpus.

Syllable type frequency tabulations for "Queen Srisawarinthira" were made on the basis of the underlying vowel lengths as indicated by Thai orthography, rather than on the basis of the phonetic transcription provided by Theraphan. Theraphan transcribes a number of underlyingly long vowels as short, reflecting the vowel reduction processes associated with secondary stress position. While the actual duration supports her transcription, I judged it better to record both the underlying length and the stress environment of each syllable. (Secondary stresses were inferred from syllable type and Theraphan's metrical marking, which does not distinguish between secondary and zero stress.)

¹⁷ Since the categorization calculations below are concerned with only one factor, vowel length, one-variable analyses of variance were used.

conditioning effect of syllable duration; under secondary stress, coda types have no significant effect on vowel duration. Differences between long and short vowels are significant under primary stress, with the exception of codaless vowels. No significant differences exist between the durations of lexically long and short vowels under secondary stress, as Siripong, Gandour, and Harper (1998) found. (Since CV? syllables are not found under secondary stress, only syllables with oral codas can be directly compared.) (“CV?, 1°” is to be read “CV? syllables under primary stress”; 0° indicates stresslessness.)

Table 1. One-variable ANOVA results for vowel length, within stress types

Syllable types	Results
(CV?, 1°) vs. (CV:, 1°)	$p > 0.05$
(CVN, 1°) vs. (CV:N, 1°)	$p < 0.001$
(CVG, 1°) vs. (CV:G, 1°)	$p < 0.05$
(CVS, 1°) vs. (CV:S, 1°)	$p < 0.05$
(CV?, 2°) vs. (CV:, 2°)	n/a
(CVN, 2°) vs. (CV:N, 2°)	$p > 0.05$
(CVG, 2°) vs. (CV:G, 2°)	$p > 0.05$
(CVS, 2°) vs. (CV:S, 2°)	$p > 0.05$

Table 2. One-variable ANOVA results for stress conditions, within syllable types

Syllable types	Results
(CV?, 1°) vs. (CV?, 2°)	n/a
(CVN, 1°) vs. (CVN, 2°)	$p < 10^{-15}$
(CVG, 1°) vs. (CVG, 2°)	$p < 10^{-9}$
(CVS, 1°) vs. (CVS, 2°)	$p < 10^{-5}$
(CV:, 1°) vs. (CV:, 2°)	$p < 10^{-20}$
(CV:, 2°) vs. (CV, 0°)	$p > 0.05$
(CV:N, 1°) vs. (CV:N, 2°)	$p < 10^{-5}$
(CV:G, 1°) vs. (CV:G, 2°)	$p < 0.01$
(CV:S, 1°) vs. (CV:S, 2°)	$p > 0.05$

Differences in the duration of vowels under primary versus secondary stress are strongly significant. The only exception is in the case of CV:S syllables, where the sample size was not large enough to support reliable statistical comparisons.

A brief reflection should make the import of these comparisons clear. Because vowel length is a *relative* contrast, when a listener is presented with some utterance to be perceived, categorized, and parsed, he or she must first make an estimation of the general rate of speech. Because vowel duration is also dependent on coda type and stress, the listener must properly perceive the following coda (if any) and gauge whether the syllable falls into the general range of durations representing primary-stressed, secondary-stressed, or stressless syllables. In calculating the probability that any particular vowel will be miscategorized by length, then, we need only compare the syllables in the coda-stress environment in which it occurs against other syllables of the same coda-stress environment. For instance, CVN syllables under primary stress need only be compared against CV:N syllables under primary stress. (The only exception to this generalization is the case of (CV:, 2°) and (CV, 0°) syllables, which do not differ significantly from each other.)

We may now proceed to calculating probabilities of miscategorization by the listener for each coda type and stress condition.

4. Tabulated data

Table 3 gives the data on which perceptual probabilities were calculated. The column labeled “N” gives the sample size for each syllable type, while column “K” gives the sample size as a proportion of syllables having the same stress and coda type. Thus, for instance, CVN syllables make up 44.5 percent ($= 117 / (117 + 146)$) of nasal-final syllables in McFarland’s list, but 58.1 percent of nasal-final syllables in Theraphan’s “Queen Srisawarithira” text. CV? and CV: syllables are treated as a single superclass for this purpose.

Table 3. Mean durations, standard deviations of durations, and sample sizes for syllable types under primary stress; data from Siripong et al. (1998)/McFarland (1942) and Theraphan (1977)

Syllable type	Siripong et al./McFarland				Theraphan			
	(Vowel interval durations)				(Syllable durations)			
	Mean duration (msec)	Std. deviation (msec)	N	<i>K</i>	Mean dur. (msec)	Std. dev. (msec)	N	<i>K</i>
CV?	(50)	(13.33)	26	0.149	254	46	5	0.079
CV:	(87)	(24.33)	148	0.851	298	88	58	0.921
CVS	50	13.33	145	0.589	251	62	44	0.595
CV:S	87	24.33	101	0.411	292	101	30	0.405
CVG	50	13.33	49	0.480	287	95	21	0.538
CV:G	87	24.33	53	0.520	364	117	18	0.462
CVN	80	6.67	117	0.445	284	65	86	0.581
CV:N	145	15	146	0.555	326	81	62	0.419

The constants in column *K* will be used as rough measures of lexical and textual frequency of occurrence to scale the probability curves for each syllable type. The data in the two *K* columns are of course the results of specific samples and show a good deal of variation. Presumably, each speaker will have his or her own notion of the relative frequency of syllable types, based on his or her own experience.

5. Calculated probabilities of miscategorization

For every pair of long-short vowel distributions, there will be some duration at which the two categories are instantiated with equal frequency. Ignoring contextual factors (which doubtless play a very large role in most real-life situations), vowels that are shorter in duration than the cut-off point will be classed by the listener as short, while vowels longer will be classed as long.

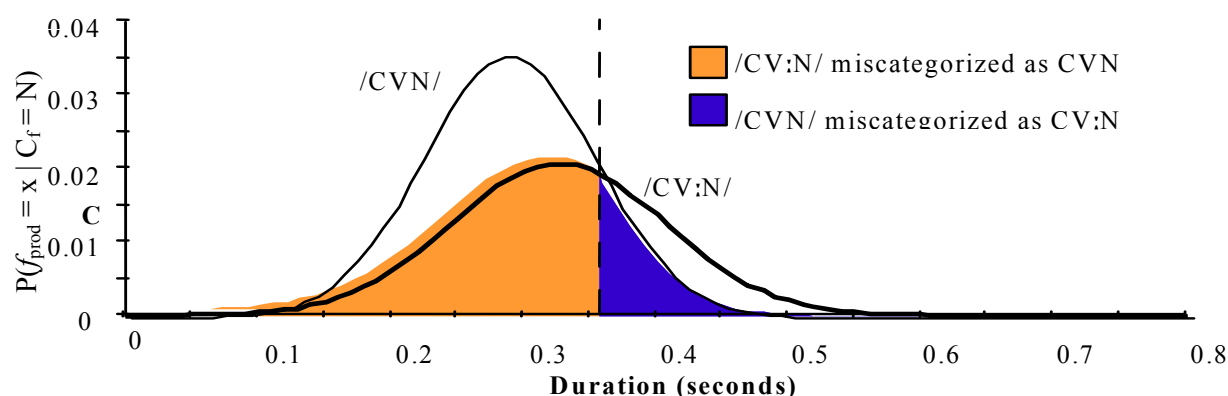
Before considering open syllables, let us take as an example nasal-final syllables in the data set from Theraphan (1977). We have already calculated *K* for CVN and CV:N, the proportion of each syllable type to all nasal-final syllables under primary stress. Given the mean and standard deviations for CVN in table 4, the expected frequency of vowels in CVN syllables having duration *x* is approximated by:

$$(15) \quad F_e(x) = K \frac{1}{s\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\bar{X}}{s}\right)^2}$$

For CVN and CV:N under primary stress, this yields figure 2. The two normal curves in figure 2 intersect at around 356 milliseconds, marked by the dashed vertical line: for durations shorter than 356 msec, CVN syllables are more plentiful, while beyond that point CV:N syllables predominate. On purely probabilistic grounds, a syllable of less than 356 milliseconds' duration is thus more likely to have been intended by its speaker to have a short vowel, while a syllable of greater than 356 msec is more likely to have been

intended to have a long vowel. A listener who perceives a syllable with a nasal coda would, other things being equal, therefore be justified in supposing that the vowel is contrastively long or short, based on whether its duration is greater or lesser than 356 msec. Naturally, other things are usually not equal: at the least, listeners must make constant adjustments for the speaker's perceived rate of speech, and they have available to them lexical, grammatical, and contextual information as they resolve raw perceived sounds into meaningful speech. But in the absence of contextual cues, and making allowance for variation in speakers' overall rates of speech, something like this probabilistic reckoning of vowel length is likely to serve listeners as an initial categorization strategy.

Figure 2. Expected portions of vowel length distributions in nasal-final syllables that are likely to be misinterpreted



The region from 0 to 356 msec in figure 3 includes roughly 86.5% of /CVN/ syllables, and 64.5% of /CV:N/ syllables. This amounts to approximately 77.3% of all nasal-final syllables, when we take into account the proportion that each syllable type represents of all nasal-final syllables. Similarly, we expect 13.5% of /CVN/ syllables and 35.5% of /CV:N/ syllables to be longer than 356 msec, or 41.9% of all nasal-final syllables. We would predict, then, that, other things being equal, our probabilistic listener will classify 77.3% of nasal-final syllables as having short vowels, and 22.7% as having long vowels. These probabilities are summarized in table 4. Note that the first three lines of table 4 give probabilities relative to *all* nasal-final syllables, not just to the intended syllable type.

Table 4. Categorization probabilities for vowel length in nasal-final syllables; data from Theraphan (1977).

$P(f_{\text{perc}} = x \mid C_f = N)$	$f_{\text{prod}} = V$	$f_{\text{prod}} = V:$	$\sum P(f_{\text{perc}} = x)$
$f_{\text{perc}} = V$	0.503	0.270	0.773
$f_{\text{perc}} = V:$	0.078	0.149	0.227
$\sum P(f_{\text{prod}} = x)$	0.581	0.419	
$\sum P(f_{\text{perc}} = y \mid f_{\text{prod}} = x)$	0.135	0.645	

The first two lines of table 4 report the probabilities that any given nasal-final syllable will have a particular intended vowel length, and a particular perceived vowel length; thus, 50.3% of nasal-final syllables ought to have been intended, and correctly perceived, as having a short vowel, 7.8% ought to have been intended with a short vowel but perceived with a long vowel, etc. The two shaded cells represent miscategorizations, where the duration of the vowel is such that the listener is likely to miscategorize the vowel for length. Column three of lines one and two, labeled ' $\sum P(f_{\text{perc}} = x)$ ', gives the

sums of probabilities for perception, while line three gives the sums of probabilities for production (note that the latter values are equal to the constant K for each category). Line four, ‘ $P(f_{\text{perc}} = y | f_{\text{prod}} = x)$ ’, reports the proportion of vowels *in each category* which are likely to be misperceived, derived by $\frac{P(f_{\text{perc}} = x | f_{\text{prod}} = x)}{P(f_{\text{prod}} = x)}$. In this case, 13.5% (= 0.078 / 0.581) of /CVN/ syllables are likely to be miscategorized as having long vowels, but 64.5% of long vowels are likely to be miscategorized as short. Note that overall, these probabilities predict that the number of vowels *perceived* as short (77.3%) should be greater than the number of vowels *produced* as short (58.1%). We may regard this as a perceptual source of pressure on the vowel length contrast in this environment.

Table 5 gives results of analogous categorization probabilities for each coda type from each data set.

Table 5. Categorization probabilities for all coda types; data from Siripong et al. (1998)/McFarland (1942) and Theraphan (1977).

$P(f_{\text{perc}} = x)$	Siripong et al. /McFarland			Theraphan		
	$f_{\text{prod}} = V$	$f_{\text{prod}} = V:$	$\sum P(f_{\text{perc}} = x)$	$f_{\text{prod}} = V$	$f_{\text{prod}} = V:$	$\sum P(f_{\text{perc}} = x)$
$C_f = \emptyset$						
$f_{\text{perc}} = V$	0.076	0.056	0.132	0.000	0.000	0.000
$f_{\text{perc}} = V:$	0.073	0.795	0.868	0.079	0.921	1.000
$\sum P(f_{\text{prod}} = x)$	0.149	0.851		0.079	0.921	
$\sum P(f_{\text{perc}} = y f_{\text{prod}} = x)$	0.491	0.066		1.000	0.000	
$C_f = N$						
$f_{\text{perc}} = V$	0.444	0.001	0.445	0.503	0.270	0.773
$f_{\text{perc}} = V:$	0.000	0.554	0.555	0.078	0.149	0.227
$\sum P(f_{\text{prod}} = x)$	0.445	0.555		0.581	0.419	
$\sum P(f_{\text{perc}} = y f_{\text{prod}} = x)$	0.001	0.002		0.135	0.645	
$C_f = G$						
$f_{\text{perc}} = V$	0.434	0.109	0.543	0.432	0.237	0.669
$f_{\text{perc}} = V:$	0.046	0.411	0.457	0.106	0.224	0.331
$\sum P(f_{\text{perc}} = x)$	0.480	0.520		0.538	0.462	
$\sum P(f_{\text{perc}} = y f_{\text{prod}} = x)$	0.096	0.210		0.197	0.514	
$C_f = S$						
$f_{\text{perc}} = V$	0.574	0.132	0.705	0.546	0.274	0.821
$f_{\text{perc}} = V:$	0.016	0.279	0.295	0.048	0.131	0.179
$\sum P(f_{\text{prod}} = x)$	0.589	0.411		0.595	0.405	
$\sum P(f_{\text{perc}} = y f_{\text{prod}} = x)$	0.027	0.321		0.081	0.676	

Three regularities evident in table 5 deserve comment. First, the probabilities for miscategorization are uniformly greater for the data from Theraphan (1977) than for the data from Siripong et al. and McFarland. Several factors might be responsible for this. The data from Theraphan (1977) is based on an orally recited text. The text is a carefully prepared one, but it seems at least likely that the overall rate of speech in the text is greater than in Siripong et al.’s constructed corpus. An overall faster rate of speech would shrink the range of duration used to express the vowel length contrast; if the variance of each syllable type shrank less than the overall duration scale, the result would be more overlap between categories and more possibility for miscategorization. Alternately, the differences may be due to

differences in sample size for each syllable type, or to the unquantified, unranked nature of McFarland's list of most frequent words.

It is notable that long vowels are consistently at a greater risk of being miscategorized than short vowels. This is true for both data sets, and for all syllable types with oral codas. The only exception is the case of syllables lacking an oral coda. In the Siripong et al./McFarland data set, roughly 48.5 percent of short vowels are likely to be miscategorized. This is a much larger figure than for any of the other syllable types, and it seems likely that the contrastive integrity of the vowel length distinction is in jeopardy in this environment. The case of codaless syllables in Theraphan's data is especially striking: here, the contrast between short and long vowels is completely obliterated, for the entire sample of CV? syllables falls within the durational range of the much more numerous CV: set. Figures 3 and 4 show the expected distributions of CV? and CV: for each data set.

Figure 3. Estimated distributions of durations for vowel intervals in CV? and CV: syllables; data from Siripong et al./McFarland

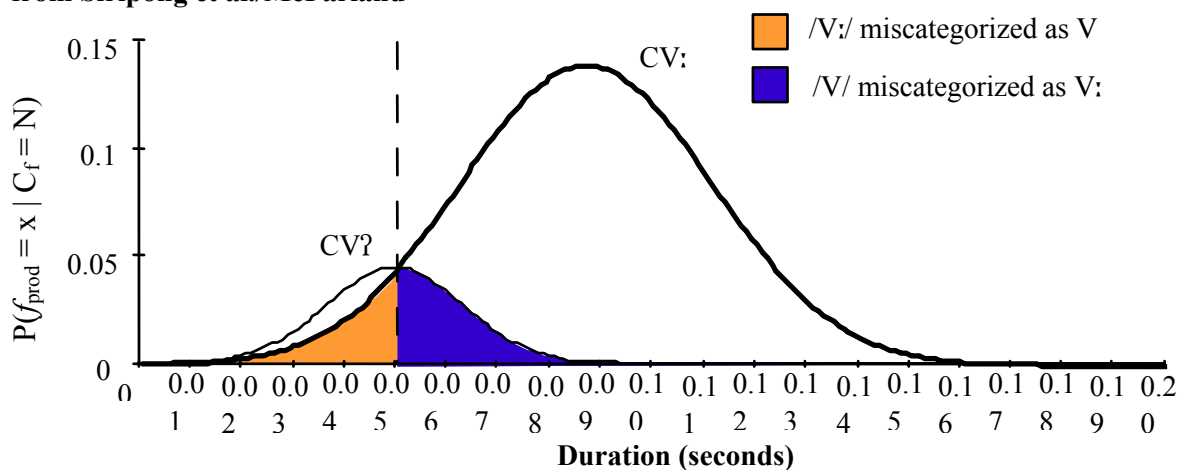
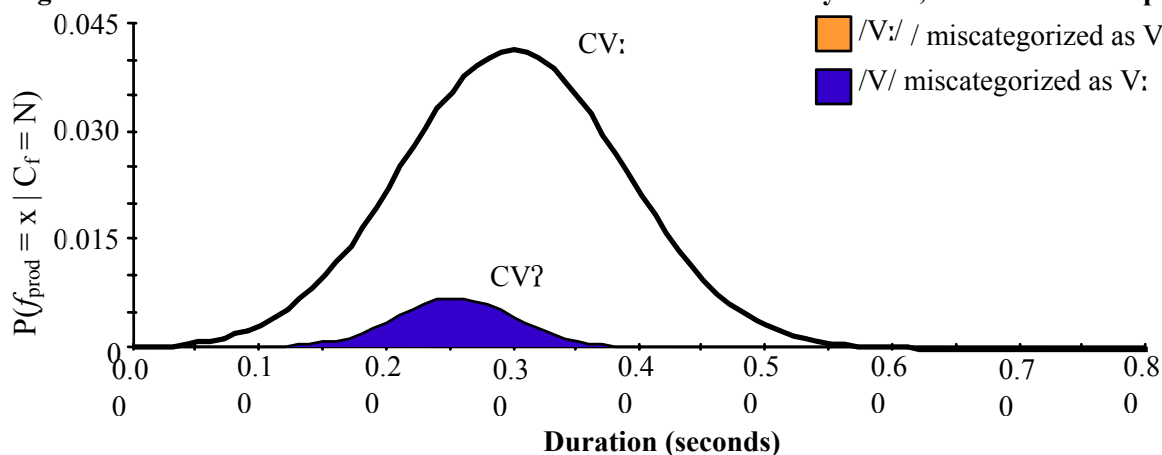


Figure 4. Estimated distributions of durations for CV? and CV: syllables; data from Theraphan



Note that these results match our requirements for confirming evidence for the displaced contrast hypothesis. The contrast between /V/ and /V:/ is in jeopardy in codaless syllables, but the threat to the

contrast lies in a particular direction, affecting short vowels more gravely than long ones. A similarly high probability of confusion is not found for other syllable types.

Under these circumstances, the addition of a glottal coda to stressed /CV/ syllables is an effective means for the speaker to signal the contrastive shortness of the vowel. A glottal stop sharply delineates the temporal range of the vowel, abruptly cutting off the vowel's formants. And since the glottal stop contributes no VC transitions of its own, it is no danger of encroaching on the contrastive territory of the oral obstruent codas [p t k].

6. Conclusions

This study found that predictable glottal codas in stressed syllables of Standard Thai occur in an environment where the vowel length contrast of the preceding vocalic nucleus is compromised by a combination of similar phonetic durations and widely disparate textual frequencies for CV? versus CV: syllables. Although the actual amount of phonetic overlap between the two syllable types varied widely in the two data sets examined, the same pattern was evident in both.

These results must be treated with some caution, because, as discussed above, it is not known for certain whether vowel quality and phonation type play any significant role in the recognition of short and long vowels in codaless syllables. However, I am cautiously optimistic that direct evidence, in the form of controlled perceptual categorization tests with native speakers of Thai, would confirm the indirect results reported here.

If glottal codas are indeed displaced markers of vowel length, we have a contrastive motivation for the epenthesis of glottal codas in loanwords and hyperarticulated forms. This allows us to resolve the difficulties in previous accounts of Thai glottal codas. A non-metrical motivation for the epenthesis allows us to dispense with the Stress-to-Weight Principle, and so to obtain a simpler analysis of the metrical system. It also means that the iambic stress pattern of so many Thai lexical items is epiphenomenal rather than explanatory: word-final CV? syllables *are* metrically heavy, but not because of a metrical constraint to that effect. Rather, they are heavy because of underlying segmental material (a long vowel or an oral coda), or because of the need to maintain the long/short vowel contrast in the stressed environment.

Appendix 1

Postscript on formalizing the analysis

I have deliberately avoided appeal to particular formalisms in this report, and I will not attempt here to develop a full proposal for the implementation of the displaced vowel length contrast in the formal grammar. However, it is appropriate to sketch the outlines of a proposal.

First, any formal account of the displaced contrast must account for both production and perception. On the production side, if glottal codas—especially glottal codas in hyparticulated speech—are epenthesized, there must be some mechanism for the epenthesis. If speakers are sensitive, even in a general way, to the likelihood of their listeners confusing their articulations, it might be reasonable to incorporate this into the constraints that are responsible for the maintenance of the long/short vowel contrast. The exact formulation of such constraints is of course a matter dependent on wider issues; but as a place of departure, we might adapt Boersma’s (1998:176) “*REPLACE” family of constraints to include the speaker’s tolerance for such ambiguity.

$$(16) \quad *REPLACE (f: x, y, p) \equiv \exists x_i \in f_{spec} \wedge \exists y_i \in f_{perc} \Rightarrow P(|x_i - y_i| > d) \leq p$$

The probability that the perceived category y on tier f is different from the specified category x by any positive distance d is not greater than p .

As we have seen, the level of tolerance can shift under differing circumstances, so that p must be regarded as a parametric rather than a categorial value.

On the perceptual side, the grammar must have some mechanism of converting the what is heard—the glottal coda, or more exactly, the sudden cessation of the vowel formants without any accompanying VC transitions—to the perceived vowel length value of shortness. Likewise, the lack of an abrupt interruption to the vowel is perceived as a long vowel length, even though the actual vowel durations may be near the periphery of the perceived category’s instantiations. Again, we might modify Boersma’s (1998:176) constraint schema:

$$(17) \quad TRANSMIT-PERC (f: x, g: y / e) \equiv \exists x_i \in f_{perc} \Rightarrow \exists y_i \in f_{spec}$$

The category x_i on a tier f in the specification corresponds to category y_i on tier g in the perceptual output in environment e .

Note that TRANSMIT-PERC maps perceived features onto specified (underlying) features, the opposite of Boersma’s TRANSMIT constraint. The variety of TRANSMIT-PERC constraint responsible for the perception of glottal codas as short vowel length would then be TRANSMIT-PERC (V length: V:, C_f: ? / stress).

Appendix 2

Kolmogorov-Smirnov goodness-of-fit test results (.05 confidence level)

Syllable type	N	D statistic	Critical value	Consistent with normal distribution?
1° stress				
CV:	58	0.183	0.179	N
CV:G	17	0.101	0.318	Y
CV:N	61	0.187	0.174	N
CV:S	29	0.178	0.270	Y
CV?	5	0.169	0.565	Y
CVG	21	0.200	0.294	Y
CVN	86	0.163	0.147	N
CVS	44	0.172	0.205	Y
2° stress				
CV:	50	0.162	0.192	Y
CV:G	3	0.385	0.708	Y
CV:N	9	0.267	0.432	Y
CV:S	1	(Insufficient data)		n/a
CV?	0	(Insufficient data)		n/a
CVG	28	0.139	0.270	Y
CVN	38	0.149	0.221	Y
CVS	12	0.167	0.375	Y
0° stress				
CV	94	0.111	0.140	Y
Aggregate statistics				
All long vowels (V:), 1° stress	93	0.250	0.141	N
All short vowels (V), 1° stress	94	0.252	0.140	N
All long vowels (V:), 2° stress	63	0.129	0.171	Y
All short vowels (V), 2° stress	78	0.122	0.154	Y
All long vowels (V:)	93	0.180	0.141	N
All short vowels (V)	94	0.250	0.140	N
All 1° stress	93	0.170	0.141	N
All 2° stress	94	0.226	0.140	N
All 0° stress	94	0.111	0.140	Y

For three of the primary-stressed syllable types, CV:, CV:N, and CVN, the value of the test statistic D lies beyond the critical value at the 0.95 confidence level. Given the relatively narrow margins by which they exceed the critical margin, however, I will treat them as if they were normal. Note that among the aggregated data, only secondary-stressed and unstressed vowels have normally distributed durations. For unstressed vowels, this is not surprising, since only a single syllable type is concerned. Normally, distributed durations for secondary-stressed vowels are also to be expected, because differences between the durations of long and short vowels in different syllable types under secondary stress are not significant (see also Siripong et al. 1998 on this).

Appendix 3

Outline of Bennett's (1995) optimality-theoretic analysis of stress placement and glottal epenthesis

In Bennett's (1995) optimality-theoretic analysis, stress-placement constraints (essentially Hayes' (1994) *End Rule-Right*), and the Stress-to-Weight Principle dominate faithfulness constraints against unsponsored moras and segmental material. (i) illustrates the constraint interaction for the normal style:

- (i) Word-final glottal epenthesis, normal style
 a. End Rule-Right, SWP >> *Insert- μ

b.

	/rajà/ 'space' (Normal style)	End Rule-R	SWP	*Ins- μ
+ (1)	[ra.'jàʔ] _o			*
(2)	[ra.'jà] _o		*!	
(3)	[ra.jà] _o	*!		

For word-medial CVʔ syllables in the isolative style, *Align-R* (σ , *PrWd*) is added to the list of undominated constraints. This constraint forces every syllable to form a prosodic word, and therefore to receive main stress.

- (ii) Word-medial glottal epenthesis in the isolative style
 a. *Align-R* (σ , *PrWd*) Align the right edge of every syllable with the right edge of a prosodic word.
 b. End Rule-Right, Al-R (σ , *PrWd*), SWP >> *Insert- μ

c.

	/sanùk/ (Isolative style)	End Rule-R Al-R (σ , <i>PrWd</i>)	SWP	*Ins- μ
+ (1)	['sàʔ] _o ['nùk] _o			*
(2)	['sa] _o ['nùk] _o		*!	
(3)	[sa.'nùk] _o	Al-R (σ , <i>PrWd</i>)!		
(4)	[sa] _o ['nùk] _o	End Rule-R!		

[ʔ] is presumably the epenthetic segment because any other coda stop would involve inserting contrastive place features, and because a lengthening of the nuclear vowel would interfere with the vowel length contrast. Bennett (1995) attempted to link the glottal character of the epenthetic stop with the observed tendency in Thai for all coda obstruents to be glottalized, at least in word or phrase final position. Whether this analysis can be maintained is doubtful, but the general glottalization of obstruent codas is noteworthy.

References

- Apiluck Tumtavitikul. 1993. Perhaps, the tones are in the consonants? *Mon-Khmer Studies* 23:11–41.
- Aporn Surinramont. 1973. Some aspects of underlying syllable structure in Thai: evidence from Khamphuan—a Thai word game. Urbana: University of Illinois M.A. thesis. (Published in *Studies in the Linguistic Sciences* 3.2:121–142.)
- Bennett, J. Fraser. 1995. Metrical foot structure in Thai and Kayah Li. Urbana: University of Illinois dissertation.
- Bennett, J. Fraser. 1998. Word stress in Standard Thai: footed or prominential? Dallas, Tex.: Summer Institute of Linguistics. ms.
- Boersma, Paul. 1998. *Functional phonology*. The Hague: Holland Academic Graphics.
- Broselow, Ellen, Su-I Chen, and Marie Huffman. 1997. Syllable weight: convergence of phonology and phonetics. *Phonology* 14:47–82.
- Duryea, David S. 1991. Issues in Thai template phonology. *Phonology at Santa Cruz*, vol. 2, ed. by Armin Mester and Scarlett Robbins, 33–58. Santa Cruz: Syntax Research Center, University of California, Santa Cruz.
- Eisner, Jason. 1997. FootForm decomposed: using primitive constraints in OT. *MIT Working Papers in Linguistics* 31.
- Haas, Mary R. 1964. *Thai–English student’s dictionary*. Stanford: Stanford University Press.
- Hayes, Bruce. 1994. *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.
- Hung, Henrietta. 1993. Iambicity, rhythm, and non-parsing. Waltham, Mass.: Brandeis University. ms.
- Henderson, Eugénie J. A. 1949. Prosodies in Siamese: a study in synthesis. *Asia Major (New Series)* 1.2:189–215.
- Kager, René. 1993. Alternatives to the Iambic–Trochaic Law. *Natural Language and Linguistic Theory* 11:381–432.
- McFarland, George. 1942. *Thai-English dictionary*. Stanford: Stanford University Press.
- McCarthy, John, and Alan Prince. 1995. Faithfulness and reduplicative identity.
- Prince, Alan. 1990. Quantitative consequences of rhythmic organization. *CLS 26: Papers from the 26th Regional Meeting of the Chicago Linguistic Society*, vol. 2: Parassession on the syllable in phonetics and phonology, ed. by Michael Ziolkowski, Manuela Noske and Karen Deaton, 357–398. Chicago: Chicago Linguistic Society.
- Prince, Alan, and Paul Smolensky. 1993. *Optimality theory*. New Brunswick, NJ and Boulder, CO: Rutgers University and University of Colorado. ms.
- Rungpat Roengpitya. 1999. A perceptual experiment on vowel length in Thai. Paper presented to the 14th International Congress of Phonetic Sciences, San Francisco, August 1–7, 1999.
- Siripong Potisuk, Jack Gandour, and Mary P. Harper. 1994. F0 correlates of stress in Thai. *Linguistics of the Tibeto-Burman Area* 17.2:1–27.
- Siripong Potisuk, Jack Gandour, and Mary P. Harper. 1998. Vowel length and stress in Thai. *Acta Linguistica Hafniensia* 30:39–62.
- Suzuki, Keiichiro. 1995. Consequences of iambic stress system in Southern Paiute. Paper presented to the annual meeting of the Linguistic Society of America, New Orleans, 6 January 1995.
- Theraphan Luangthongkum. 1977. *Rhythm in Standard Thai*. Edinburgh: University of Edinburgh dissertation.
- Udom Warothammasikkhadit. 1967. Some phonological rules of Thai. *Journal of the American Orientalist Society* 87:541–574.
- Walker, Rachel. 1997. Prominence-driven stress. Santa Cruz: University of California-Santa Cruz MS (distributed by Rutgers Optimality Archive, ROA 172-0197).