

# Is there a prosodic word in Vietnamese?\*

Andrea Pham  
University of Florida

This paper examines clitics in colloquial Vietnamese to show that there is a possible prosodic structure at the word level in the language and that the domain of this structure does not correspond to a syntactic constituent. A clitic exhibits a phonological dependency on the material to its left. The tone of the clitic remains while all segments are deleted. The tone is then realized on either a syllabic nasal or as a lengthened part of a vowel or a glide depending on the syllabic structure of the host and of the clitic itself. If the host ends in an obstruent, the clitic surfaces as a homorganic syllabic nasal. If the host ends in a vowel or a glide, the clitic surfaces as a lengthened part of that vowel or glide, unless it has a rhyml nasal, in which case it surfaces with its own nasal.

## 1. Introduction

In English, many monosyllabic function words appear as either strong forms or weak forms, e.g., ‘at’ in ‘look at’ versus ‘at’ in ‘at home’ (Selkirk 1996). A similar phenomenon occurs in colloquial Vietnamese. For example, *ở* ‘to live, stay’ in *Anh ấy ở đâu?* ‘Where does he stay?’ appears in a strong form while *ở* ‘in, at’ in *Anh ấy ở đâu?* ‘Where is he?’ appears in a weak form (further examples in Hoàng & Hoàng 1975:96).

However, certain words are often further reduced, so that the whole syllable is deleted except for the tone. In the following examples, the clitic and the host are given in parentheses and a morpheme-by-morpheme translation is provided together with an idiomatic translation. Tones are referred to by use of a number that follows the syllable, e.g., [ba1] has tone 1, [ba2] has tone 2, etc. Characteristics of tones are described in Table 4.

- (1) a. (biɣt7 ba:w1).niɣw1 xoŋ1 → (biɣt7 ɲ1) niɣw1 xoŋ1  
know how much inter part  
‘(Do you) know how much (it is)?’

---

\* I would like to thank Marc Brunelle for very useful feedback. All errors are mine.

- b. nɔ3 (la2 ka:j3) zi2 kuɤ5 ma:j2 → nɔ3 (la2 a3) zi2 kuɤ5 ma:j2  
 3sg be cl what of 2sg  
 ‘Who is he/she to you?’

In (1a) the host [biɤt7] ends with a final stop. The clitic [ba:wɪ] attaches leftward, its segments are deleted and its tone is realized on a homorganic nasal which assimilates to the place of articulation of the final stop [t] in the host. In (1b) the host [la2] is an open syllable and the clitic [ka:j3] attaches leftward. All segments of [ka:j3] are deleted except the tone. The vowel of the host is lengthened to carry the ‘floating tone’ of the clitic.

This paper will show that a clitic is a phonological subordination of a neighboring word. In this process a word is cliticized onto an adjacent word to the left. All segments are deleted, but the tone remains. This tone is realized on either a syllabic nasal or a vowel depending on the phonetic characteristics of the segment at the edge of the host, i.e., the segment to which the clitic attaches.

The term ‘clitic’ is therefore used in this paper to refer to function words such as those shown in (1) that surface as ‘reduced’ forms in colloquial Vietnamese. The reduction of these forms is not just a random process but is consistently determined by the structure of the rhyme of the host, and of the clitic itself. This phenomenon has received very little attention in the Vietnamese literature. Few such forms are mentioned in discussion of stress and the morphosyntactic properties of Vietnamese. The paper will provide a detailed description of Vietnamese clitics together with enough data so that interested researchers will be able to work on prosodic domains in Vietnamese and perhaps even on typological issues in the theory of prosodic structure. It also provides some evidence for a possible prosodic structure in Vietnamese. However, it does not attempt to account for stress patterns nor make claims as to the nature of the prosodic hierarchy of the language.

The paper is organized as follows. Section 2 provides relevant background on the syllable structure and prosodic characteristics of Vietnamese. Section 3 describes the properties of clitics, examines the processes of cliticization, and uses spectrographic representations to show that clitics are indeed ‘reduced’ forms of their full counterparts and that the duration of clitics is not always shorter than that of their full counterparts. Section 4 discusses a small group of superficially similar contracted forms in Southern dialects to show that this contraction process is not cliticization. The conclusion is in section 5.

The data are from daily conversations recorded in Vietnam and from recordings of participants in interview’s telecast on Saigon City Television and broadcast on the Vietnamese Radio Station in Toronto. The spectrograms come from sample recordings of the author’s speech.

## 2. Background

### 2.1 Syllable structure

Vietnamese is mainly monosyllabic, i.e., one syllable of C(w)V(C) syllable structure can form a word. The only consonant cluster is Cw. The C, V, and tone are mandatory; the rest is optional. V can be either a short or a long vowel, or a diphthong. Compound words and reduplicative forms have more than one syllable. This paper

discusses the Hanoi or standard dialect. Table 1 shows the phonemic initial consonant inventory.

	labial	coronal	palatal	dorsal	glottal
voiceless stop		t	c	k	ʔ
voiced stop		d			
aspirated stop		t <sup>h</sup>			
voiceless fricative	f	s		x	h
voiced fricative	v	z		ɣ	
lateral approximant		l			
nasal	m	n	ɲ	ŋ	

Table 1. Phonemic initial consonants in the Hanoi dialect

Table 2 shows the phonemic vowel inventory of the Hanoi dialect. There are eleven vowels and three diphthongs. Two vowels /a/ and /ə:/ have a length contrast. Back vowels are rounded. Short vowels occur only in closed syllables, i.e., syllables that end with a consonant or a glide.

	front	central	back
high	i	ɨ	u
mid	e	ə ə:	o
low	ɛ	a a:	ɔ
diphthong	iə	ɨə	uə

Table 2. Phonemic vowel inventory in the Hanoi dialect

Only a limited numbers of consonants and glides occur finally in syllables: /p, t, k, m, n, ŋ, w, j/. As we will see later, assimilation of cliticization in Vietnamese is a phonetic process; therefore, certain final allophones are important. In Table 3, the palatals and the plain and labialized dorsals are allophones of the same phoneme, i.e., [c], [k] and [k<sup>p</sup>] are allophones of /k/ and [ɲ], [ŋ] and [ŋ<sup>m</sup>] are allophones of /ŋ/.

consonants	p	t	c	k	k <sup>p</sup>
	m	n	ɲ	ŋ	ŋ <sup>m</sup>
glides	w	j			

Table 3. Phonetic final inventory in the Hanoi dialect

There are eight tones in the dialect. Tones 1 and 2 are level, tone 3 rises gradually, tone 4 is falling with either a glottal stop or creakiness, tone 5 falls then rises, and tone 6 is also falling-rising but is broken by a glottal stop. Tone 7 is rising and short. Tone 8 is falling and short. The last two tones occur only in obstruent-finally syllables.<sup>1</sup> Table 4 shows the tonal inventory. Glottalization is distinctive in tones 4 and 6. Each tone is also shown with its Vietnamese name.

tone 1 <i>ngang</i> (high, level)	tone 3 <i>sắc1</i> (rising)	tone 5 <i>hỏi</i> (falling- rising)	tone 7 <i>sắc2</i> (rising, short)
tone 2 <i>huyền</i> (low, level)	tone 4 <i>nặng1</i> (falling, glottalized)	tone 6 <i>ngã</i> (falling- rising, glottalized)	tone 8 <i>nặng2</i> (falling, short)

Table 4. Tonal inventory

## 2.2 Stress in Vietnamese: how significant is it?

The literature on clitics often discusses this category within a prosodic structure whose hierarchy proceeds from the lowest level to the highest level as follows: syllable, foot, prosodic word, phonological phrase, intonational phrase and utterance (Selkirk 1978), e.g. “Prosodic clitic is definable with respect to the category *prosodic word*” (Selkirk 1978, 1996: 441). Studies of on Vietnamese focus mainly on the level of syllable and any possible higher prosodic level has received little attention. The only study that has so far addressed this issue is Hildebrandt et al (2006) and the claim there is that the prosodic word does not exist in Vietnamese. Consequently, the above prosodic hierarchy is not universal. However, the evidence presented in this paper will show that the prosodic word almost certainly exists in Vietnamese.

‘Foot’ is a construct that is necessarily based on stress (Liberman & Prince 1977); ‘prosodic word’ dominates at least one foot (Peperkamp 1997: 16). Clitics are also defined as deficient items with the deficiency largely attributable to lack of stress. It is generally assumed that stress in Vietnamese is not crucial and that if there is any, it is secondary to tone (Ngô 1984).

This section discusses stress in Vietnamese and focuses on those aspects that are relevant to clitics. Stress does play a significant role in distinguishing morphosyntactic structures such as certain compounding and reduplicative patterns/types, and lexical and functional words (Hoàng & Hoàng 1975, Ngô 1984, Cao 1978 & 1998). Vietnamese clitics, just as in other languages, are ‘little words,’ i.e., reduced forms, that always lack stress.

<sup>1</sup> There is also a six-tone view. This paper adopts an eight tone system. See details and arguments for each view in Pham 2003.

The general assumption in the literature on Vietnamese is that because it is monosyllabic and has lexical tones, stress plays no role in the phonology of the language; it may not even exist. For example, Honey (1989: 241) states: ‘Being monosyllabic and tonal, Vietnamese is unable to express stress by tone.’ Perhaps that is the reason why clitics have been overlooked. In one view, every syllable is said to be pronounced with the same energy, i.e., with similar intensity and duration (Emeneau 1951, Benedict 1948, Hồ 1976, Ngô 1999). This characteristic is claimed to be a major obstacle for English speakers to overcome in learning Vietnamese (Ngô 1999) and, vice versa, to be a source of difficulty for Vietnamese speakers in learning English since their ‘first inclination is to give full stress to each syllable of a polysyllabic word (Honey 1989).

Other researchers claim that in Vietnamese there is no distinction between stressed and unstressed syllables since each syllable has some degree of stress, either primary or secondary (Lê 1976) or even three degrees of stress: heavy, medium, and weak (Thompson 1965). These authors focus on phrasal stress and different tempos of speech. Nguyễn 1990 mentions ‘main stress’ used apparently for emphatic or focused purposes, i.e., on the interrogative substantive in content questions such as *ai?* ‘who?’, *gì?* ‘what?’, *nào?* ‘which?’, and *đâu?* ‘where’ when these function as indefinite pronouns, ‘especially in women’s speech [when] they can even have negative indefinite function, provided the interrogative substantive receives very heavy stress,’ e.g. *ai* ‘who’ in *ai nói* ‘no one spoke’ (1990: 788).

Yet another view is that we find certain stress patterns in disyllabic words depending on their syntactic relationships (Cao 1978 & 1998, Nguyễn 1994). For example, in *cà chua* ‘tomatoes,’ literally ‘eggplant + sour,’ or in *chuồn chuồn* ‘dragon fly,’ the second element bears stress; therefore, these syllables are pronounced continuously, i.e., there is no pause between the two syllables. In *tóc đen* ‘black hair,’ literally ‘hair’ + ‘black,’ both syllables bear stress; consequently, they are pronounced separately. The stress pattern in certain disyllabic and polysyllabic forms and in the traditional *lục bát*, a ‘six-eight’ verse form, also shows that Vietnamese appears to be an iambic language, i.e., each foot consists of a weak syllable followed by a strong one (Pham, ms). Moreover, this view claims that stress does affect tones. For example, in unstressed syllables tone 3 (rising) is lower due to the loss of the second part, tone 5 (falling-rising) loses its rising part, and tone 4 and tone 6 are sometimes neutralized to tone 2 because they lose the second part after the glottal stop (Cao 1978).

Hoàng & Hoàng 1975 and Ngô 1984 mention stress in monosyllabic words. They connect it directly to a ‘reduction’ in intensity and duration. Hoàng & Hoàng 1975 also claim that function words are always unaccented. ‘Accented’ is defined as ‘certain syllables [being] pronounced stronger and longer than others’ (1975:68). Consequently, the absence of accent in words in this category is due to their inability to form an independent syntactic phrase, nor can they be placed at the end of a phrase. They are all proclitics except for unaccented particles, such as *à, ư, nhỉ, nhé, mà, cơ, sao, không, chứ,* and *à* that end a phrase (1975: 94). This unaccented property makes this word category subject to significant modifications including reduction in intensity and duration (“Dans les syllabes non accentuées, c’est-à-dire celles qui se trouvent à l’intérieur d’un syntagme, il est sujet à des modifications considérables en fonction de la réduction de la durée et de l’intensité syllabique”). The accent is also said to be important in

distinguishing between the meanings of homonyms such as the indefinite article *một* and the numeral *một* ‘one.’ The former is always unaccented and the latter is always accented.

This paper assumes that every monosyllabic content word bears some degree of stress and that function words are often unstressed and can appear either in strong or weak forms depending on the focus of the utterance. The following section will show that clitics are function words and always appear as reduced forms; therefore, as weak forms they cannot bear stress.

### 2.3 What is a clitic in Vietnamese?

Cliticization occurs frequently in colloquial Vietnamese. The phenomenon is mentioned in both Hoàng & Hoàng 1975 and Ngo 1984, who says that ‘In a weekly stressed syllable, the segmental syllable is reduced or vanished in allegro speech’ (112). However, these author so not identify such reduced forms as ‘clitics’ nor do they examine how they are reduced or vanished thus revealing their unique properties. This paper will describe these properties and show that these ‘weakly stressed’ syllables are neither necessarily shorter than their full-form counterparts nor occur only in fast speech as is often claimed. They are found in informal, casual styles of speech regardless of how ‘fast’ or ‘slow’ that speech is.

Modern syntacticians usually treat clitics morphosyntactically (see Anderson 2005 for a comprehensive review). However, for Wackernagel (1892) and other traditional grammarians and Indo-European linguists, a clitic was a ‘stressless *little* word that lacks independent accent, and depends prosodically on an adjacent word” (Anderson 2005: 2). According to Selkirk (1996: 440), a prosodic clitic is ‘a morphosyntactic word which is not itself a prosodic word.’ A Vietnamese clitic attaches freely to any host on its left side. Vietnamese does have a very limited number of so-called ‘affixes’ (Hoàng and Hoàng 1975, Ngô 1984) that function like ‘words,’ i. e., they can also occur alone. Examples of some prefixes include the classifiers used to indicate animate/inanimate, such as *cái* (objects), *con* (animals), and *người* (people), or to indicate an abstract noun such as *sự*, or *việc*. Suffixes include those used to turn a noun into a verb, such as *hóa* ‘-ize’ (e.g. *kỹ nghệ hóa* ‘industrialize,’ and *lão hoá* ‘become old’), or to indicate people, such as *sĩ*, or *sur* (e.g., *họa sĩ* ‘painter,’ *ca sĩ* ‘singer,’ *nghệ sĩ* ‘artist,’ *giáo sĩ* ‘missionary,’ *giáo sư* ‘professor,’ *võ sư* ‘martial art teacher,’ *gia sư* ‘home tutor’) (see Ngô 1984 for details). However, only one or two of these behave like clitics.

Hoàng & Hoàng (1975: 94) and Ngô 1984 are also the only sources that describe the phenomenon of clitics in some detail, but they cite very few words. Hoàng & Hoàng point out that the adverb *cũng* [kũŋ<sup>m</sup>6] ‘also’ is often pronounced as [ŋ<sup>m</sup>6] and that the indefinite article *một* [mot4] is often pronounced as [m̩4]<sup>2</sup>. They represent these reduced forms with a syllabic nasal and the original tone of the clitic and state that while these words are reduced to one sonorant, they nevertheless retain a melodic pattern that is recognizable through the curve of intensity that clearly indicates a syllable in the utterance. However, these two words are the only examples they cite and they offer no further comment on, or description of, the processes. A second case of reduction they

<sup>2</sup> The numbers used in Hoàng & Hoàng 1975 for tonal reference are different from those found in this paper, although those tones are the same. These authors assume a six-tone system.

mention is that of *cái* /ka:j3/ (classifier) being reduced to [ki5] or [kɤ5] and of the conjunctions *mới* /mɤ:j3/ ‘with’ and *với* /vɤ:j3/ ‘with’ being reduced to [mi3].

Ngô (1984:90) also describes certain function words that lose segments in fast speech because of lack of stress. He says that the rhyme is lost first, e.g., *cái gì* [ka:j3 zi2] ‘what’ becomes (k3 zi2), *ai mà* [a:j1 ma2] ‘who that’ becomes (a:j1 m2), and *ngủ một tí* [ɲi5 mot8 ti3] ‘rest a while!’ becomes [(ɲi5 m8) ti3]; or the whole syllable is lost, but not the tone, e.g., (a:j1 2) from [a:j1 ma2] and (ɲim5 8) from [ɲi5 mot8]. He does show the surviving consonant to be syllabic, nor is it clear whether there is a stray tone in (ɲim5 8). He also mentions that most function words are weakly stressed and reduced within the phrase in which they occur. However, he presents the reduced forms as if they occur context-free, e.g., *con* [kɔn1] ‘classifier’ becomes [kɤ1], *cái* [ka:j3] becomes [k3], and *mà* [ma2] becomes [m2] (1984: 100). Function words can be so reduced independent of context, but they are not clitics since they do not associate with a host and their surfaces are not structurally affected by the host; therefore, they are outside the focus of this paper. The focus is on ‘true’ clitics whose surfaces depend on the structure of the hosts, i.e., a word is reduced to a single sonorant but retains its tone, as when *mang cái gì* [(ma:ɲ1 ka:j3) zi2] ‘what to bring’ becomes [(ma:ɲ1 ɤ3) zi2].

Furthermore, the paper will show that the reduction of segments in *cũng* [kuŋ<sup>m</sup>6] ‘also’ to [ɲ<sup>m</sup>6] and in *một* [mot8] to [m8] is the result of different processes: *cũng* and *một* can have different surface forms that depend on the hosts to which they attach. This paper will also show that in the case of *một* the tone is not left to stray as in Ngô 1984 says nor does the clitic surface with its original tone, as both Ngô and Hoàng & Hoàng claim.

It is now possible to propose a general definition of clitic in Vietnamese. A clitic is an unstressed function word that always appears as a reduced form and associates with its host on the left. The deleted syllable leaves behind its tone, which is realized on either a lengthened vowel/glide or as a syllabic nasal in a structural relationship between the host and the clitic.

### 3. Clitics: properties and processes

This section describes the morphosyntactic and phonological characteristics of clitics together with various processes involved in Vietnamese cliticization. In this process a word is cliticized onto an adjacent word to the left. All segments are deleted but the tone remains. This tone is realized on either a syllabic nasal or a vowel depending on the phonetic characteristics of the segment at the edge of the host.

#### 3.1 Morphosyntactic characteristics of clitics

Vietnamese clitics are function words. They can be demonstrative pronouns, adverbs, personal pronouns, negatives, conjunctions, prepositions, complementizers, quantifiers, or classifiers. The most common clitics include *cũng* ‘also,’ *có* ‘to have’ (auxiliary verb), *là* ‘to be’ (auxiliary), *mà* ‘that’ (pronoun), *ấy* ‘that’ (demonstrative), *làm sao* ‘how,’ *cái* (classifier for objects), *đấy* (emphasis particle), *đi* (emphasis

particle)<sup>3</sup>, *trong* ‘in,’ *với* ‘with,’ *bao nhiêu* ‘how much, how many,’ *một* ‘a (indefinite determiner),’ and *thì* ‘then’ (pronoun). (2) gives some examples. The host and clitic in each phrase are bracketed in both left and right columns. The left column provides the full forms of all words. The right column shows the surface forms of clitics following the cliticization process.

- (2) a.  $\varepsilon m1$  ( $c\alpha1$   $n\alpha3$ )  $\eta i5$   $z\alpha j2$  →  $\varepsilon m1$  ( $c\alpha1$   $\alpha3$ )  $\eta i5$   $z\alpha j2$   
 1sg allow 3sg rest already  
 ‘I already fired him.’
- b. ( $tuj4$   $?xj3$ )  $law3$   $lam3$  → ( $tuj4$   $j3$ )  $law3$   $lam3$   
 group that smart very  
 ‘They are very smart.’
- c. ( $dij2$   $k\alpha3$ )  $la:m2$   $t^hxt8$  → ( $dij2$   $\eta3$ )  $la:m2$   $t^hxt8$   
 not have do real  
 ‘Do not actually do it.’
- d.  $ci3$   $vij1$  ( $kew1$   $la2$ )  $mxt8$  →  $ci3$   $vij1$  ( $kew1$   $w2$ )  $mxt7$   
 sister Vinh say that lose  
 ‘Sister Vinh said (it) was lost’
- e.  $?ap1$  ( $dij4$   $la:m2$ )  $sa:w1$  →  $?ap1$  ( $dij4$   $\eta2$ )  $sa:w1$   
 2sg. plan how  
 ‘What is your plan (to deal with it)?’

Note that in (2e) the clitic is [ $la:m2$ ] and it is the first element of the adverb [ $la:m$   $sa:w$ ] ‘how.’ If [ $la:m2$ ] is a verb meaning ‘to do,’ it cannot become a clitic regardless of how colloquial or fast the speech is, as (3) shows.

- (3)  $?ap1$   $dij4$   $la:m2$   $zi2$  →  $?ap1$   $*(dij4$   $\eta2)$   $zi2$   
 2sg. plan do what  
 ‘What do you intend to do?’

The morphosyntactic constituent does not seem to play a role in the operation of cliticization. (2) shows that a clitic can attach to either a lexical word such as the verb / $c\alpha$ /, as in (2a), or a function word such as the negative / $dij$ /, as in (2c). All hosts in (2) are monosyllabic words. However, a clitic can ‘break’ a grammatical word that consists of two elements, such as a compound, a proper name, or a reduplicative form. For example, in (4a) the clitic *đó* / $d\alpha$ / attaches to the second syllable of a proper (first) name *Minh Anh*. Moreover, the clitic itself can be a disyllabic word but only the relevant syllable attaches to the host and undergoes the process. For example, in (4b) only the first syllable in the clitic *bao nhiêu* / $ba:w$   $\eta i\alpha w$ / (how much?) attaches to the host *hết* / $het$ / on its left side. In (4c) both the host, a reduplicative *lâm cẩm* / $l\alpha m$   $k\alpha m$ / ‘silly’ and the clitic

<sup>3</sup> There are many elements in Vietnamese labeled ‘particles’ in the traditional literature. Zwicky 1985 argues that ‘particles’ are not a word class and that some are actually clitics. However, this issue is not a concern of this paper.



*làm sao* /la:m sa:w/ ‘how’ are disyllabic. The clitic attaches leftward and only the first syllable *làm* attaches to the second syllable of the host.

- (4) a.  $\gamma\text{ap}^8 \text{ m}\text{in}^1 (\text{?a}\text{n}^1 \text{ d}\text{o}^3) \text{ n}\text{x}^4 \rightarrow \gamma\text{ap}^8 \text{ m}\text{in}^1 (\text{?a}\text{n}^1 \text{ n}^3) \text{ n}\text{x}^4$   
 meet Minh Anh emp emp  
 ‘(I) met Minh Anh.’  
 b.  $\text{m}\text{u}\text{x}^1 (\text{het}^7 \text{ ba:w}^1) \text{ n}\text{i}\text{xw}^1 \rightarrow \text{m}\text{u}\text{x}^1 (\text{het}^7 \text{ n}^1) \text{ n}\text{i}\text{xw}^1$   
 buy all how much  
 ‘How much was spent (on it)?’  
 c.  $\text{m}\text{in}^2 \text{ l}\text{a}\text{m}^5 (\text{k}\text{a}\text{m}^6 \text{ la:m}^2) \text{ sa:w}^1 \rightarrow \text{m}\text{in}^2 \text{ l}\text{a}\text{m}^5 (\text{k}\text{a}\text{m}^6 \text{ m}^2) \text{ sa:w}^1$   
 1sg silly how  
 ‘How silly I am!’

Because the majority words in Vietnamese are monosyllabic, many prosodic words contain only one syllable. The fact that a clitic can cliticize onto and be affected by just one syllable of a disyllabic word supports the claim that the host is not larger than a prosodic word (Selkirk 1996, Peperkamp 1997, van der Leeuw 1997).

Clitics are also insensitive to syntactic structure. In (5a) the classifier *cái* /ka:j/ ‘a’ attaches leftward to a host *năm* /nəm/ ‘five’ within the same quantifier phrase, i.e., [nəm [ka:j ba:n]<sub>NP</sub>]<sub>QP</sub> ‘five tables,’ or the clitic emerges from its phrase and attaches to a verb, i.e., [xiəŋ [ka:j ba:n]<sub>NP</sub>]<sub>VP</sub> ‘to carry a table,’ as in (5b).

- (5) a.  $\text{x}\text{i}\text{ə}\text{n}^1 (\text{n}\text{ə}\text{m}^1 \text{ ka:j}^3) \text{ ba:n}^2 \rightarrow \text{x}\text{i}\text{ə}\text{n}^1 (\text{n}\text{ə}\text{m}^1 \text{ m}^3) \text{ ba:n}^2$   
 carry five cl table  
 ‘to carry five tables’  
 b.  $(\text{x}\text{i}\text{ə}\text{n}^1 \text{ ka:j}^3) \text{ ba:n}^2 \rightarrow (\text{x}\text{i}\text{ə}\text{n}^1 \text{ n}^3) \text{ ba:n}^2$   
 carry cl table  
 ‘to carry a table’

This last point suggests that the domain of cliticization is a prosodic one. In other words, the domain of a clitic and its host does not correspond to a syntactic constituent. In this respect, Vietnamese clitics are similar to final particles in Shanghai Chinese, which attach to whatever lexical items happen to precede them in syntactic surface structure” (Selkirk & Shen 1990: 318).

However, there is one instance in which only a specific element in the phrase can cliticize. In Vietnamese a modifier follows what it modifies. The basic word order of a noun phrase is: indefinite quantifier – classifier – noun – adjective, e.g., *một cái bàn nhỏ* ‘a small table.’ When a phrase contains both a quantifier and a classifier and one of them is a clitic, it must be the quantifier not the classifier. For example, in (6a) *một* /mot/ can attach to the verb *có* /kə/ ‘to have’ on the left, but the classifier *cái* /ka:j/ cannot attach to *một*, as (6b) shows. Note that the clitic [mot] bears tone 8, but after cliticization tone 2

surfaces instead on the lengthened vowel in (6). This is because tones 7 and 8 occur only in stop-final syllables.<sup>4</sup>

- (6) a. nɔ3 (kɔ3 mot8) ka:j3 ɲa2 tɔ1 lam3 → nɔ3 (kɔ3 ɔ2) ka:j3 ɲa2  
 tɔ1 lam3  
 3sg have a cl house big very  
 ‘He/she has a very big house.’  
 b. nɔ3 kɔ3 (mot8 ka:j3) ɲa2 tɔ1 lam3 → \*nɔ3 kɔ3 (mot8 ɲ3) ɲa2  
 tɔ1 lam3

Note also that in (6a) the indefinite quantifier *môt* could cliticize leftward to attach to /ka:j/ and surface as a syllabic velar nasal. Normally clitics attach leftward. In this case there is a slight pause before the indefinite quantifier. This restriction is not seen elsewhere. The direction of cliticization will be discussed further in the following section. In general, however, clitics are free to attach to any host on the left regardless of whether it is or is not in the same morpho-syntactic domain as the host.

### 3.2 Processes of cliticization

This section describes how a clitic surfaces in different environments after all its segments are deleted. In all cases, the tone of the clitic remains and attaches leftward to either an obstruent or a vowel/glide. If the host ends in an obstruent, the clitic surfaces as a homorganic nasal. If the host ends in a vowel or a glide, the vowel/glide is lengthened and the tone of the clitic is realized on the lengthened part, unless the clitic itself contains a nasal. In that case the clitic surfaces with its original nasal and tone.

#### 3.2.1 Clitics that surface as homorganic nasals or vowels/glides

This type was seen in (2a) where the pronoun *nó* /nɔ/ ‘he/she’ cliticizes leftward and the host ends in a vowel that is lengthened to carry the tone of the clitic: (cɔ1 nɔ3) → (cɔ1 ɔ3). In (2b) the demonstrative *áy* /ɤj/ ‘that’ cliticizes leftward and the host ends in /j/, which is lengthened to carry the tone of the clitic: (tɤj4 ɤj3) → (tɤj4 j3). (The spectrograms in section 3.4 will show this.) In these two examples the clitic happens to contain the same vowel or glide as the host. However, (7c) and (7e) show that this is only a coincidence. More examples of this type are given in (7). The clitics surface as homorganic nasals in (7a-c) and as vowels/glides in (7d-g). The clitics surface as lengthened parts of vowels, not as vowels with glottal stops between the host and the clitic. (7c) and (7d) also contain two clitics. In (7c) the second clitic in italics *kuy<sup>m</sup>6* is special in that it does not have a host. This type of clitic will be discussed later.

<sup>4</sup> All clitics keep their original tones, except in cases like this. The six-tone view, which regards tones 7 and 8 as allotones of tones 3 and 4, would have to say that a tonal change accounts for this. However, in the eight-tone view, which treats two tones in stop-final syllables as separate tones, there is no change here: the clitic [mot8] surfaces with tone 2 simply because the clitic appears as a vowel and, therefore, cannot bear tone 8. Tone 2 is phonetically closest to tone 8; consequently, it replaces tone 8. [mot] is the only clitic with tone 8.

- (7) a. nɔj1 (cɤm4 kuŋ<sup>m</sup>6) diɤk8 → nɔj1 (cɤm4 m̃6) diɤk8  
 speak slow also fine  
 ‘It is okay to speak slowly.’
- b. (diŋ2 kɔ3) la:m2 t<sup>h</sup>ət8 → (diŋ2 ŋ3) la:m2 t<sup>h</sup>ət8  
 not have do real  
 ‘Do not really do it.’
- c. (viət7 təm2) bəj4 kuŋ<sup>m</sup>6 diək8 → (viət7 ŋ2) bəj4 ŋ<sup>m</sup>6 diək8  
 write nonsense also okay  
 ‘It’s ok to write nonsense (here).’
- d. kɔn1 (mɛ:w2 ma2) (bat7 kɔn1) cuət8 → kɔn1 (mɛ:w2 w2) (bat7 ŋ1) cuət8  
 cl cat comp catch Cl mouse  
 ‘The cat that caught the mouse.’
- e. ti2 (naj6 təj3) zə2 → ti2 (naj6 j3) zə2  
 since then until now  
 ‘since then until now’
- f. zu2 (sa:w1 di1) niɤ6 → zu2 (sa:w1 w1) niɤ6  
 regardless of how still  
 ‘anyhow’
- g. za1 (dɔ3 ma2) ŋɛ1 → za1 (dɔ3 ɔ2) ŋɛ1  
 go out there comp listen  
 ‘Go out there to hear (that)!’

The surface forms of clitics result from process of phonetic assimilation. For example, in Vietnamese final palatal consonants are allophones of velars; however, if the host ends with a palatal, the clitic surfaces as a palatal nasal, e.g., (diŋ4 ŋ2) in (2e) or (aŋ1 ŋ3) in (4a). Likewise, velar consonants after back rounded vowels are labialized and surface as double articulations. If the host ends in a labialized velar, the clitic surfaces as a homorganic labialized velar, as with the first clitic *trong* [cɔŋ<sup>m</sup>1] ‘in’ (8a) below. Note that this sentence also has two clitics. Although one might suspect that the surface form of the clitic, [cɔŋ<sup>m</sup>1], is indeed its original labialized nasal, this is not the case (8c) shows. Here the original labialized nasal of the clitic is deleted and the clitic assimilates to [t] in the host [ha:t] and surfaces as a coronal nasal.

- (8) a. nɔ3 (xɔk<sup>p</sup>7 cɔŋ<sup>m</sup>1) (fɔŋ<sup>m</sup>2 dɤj3) → nɔ3 (xɔk<sup>p</sup>7 ŋ<sup>m</sup>1) (fɔŋ<sup>m</sup>2 m̃3)  
 3sg cry in room emp  
 ‘He/she is crying in the room.’
- b. xi1 na:w2 (sɔŋ<sup>m</sup>1 t<sup>h</sup>i2) dɛm1 la:j4 → xi1 na:w2 (sɔŋ<sup>m</sup>1 m̃2) dɛm1 la:j4  
 when finish then bring back  
 ‘Bring (it) back when (you are) done.’

- c. nɔ3 (ha:t7 cɔŋ<sup>m</sup>1) (fɔŋ<sup>m</sup>2 dɤj3) → nɔ3 (ha:t7 ɲ1) (fɔŋ<sup>m</sup>2 ɲ3)  
 3sg sing in room emp  
 ‘He/she is singing in the room.’

### 3.2.2 Clitics without hosts

In all examples given above, the clitics lose all their segments in cliticization. However, there are some clitics that retain at least one original segment, the final nasal. A clitic keeps its final nasal if the host ends in a vowel or a glide. The clitics in (9) have hosts that do not end with obstruents and [ɛm1], [kuŋ<sup>m</sup>6], [aŋ1], and [ben1] all surface with final nasals.

- (9) a. (tʰɤj3 ʔɛm1) xɔk<sup>p</sup>7 → (tʰɤj3 ɲ1) xɔk<sup>p</sup>7 /\*(tʰɤj3 j1)  
 see 2sg younger cry  
 ‘(I) saw you cried.’  
 b. (toj1 kuŋ<sup>m</sup>6) ɲa:j3 → (toj1 ŋ<sup>m</sup>6) ɲa:j3 /\*(toj1 j6)  
 1sg also imitate  
 ‘I also imitated (it).’  
 c. (tʰi2 ʔaŋ1) mɤ:j2 → (tʰi2 ɲ1) mɤ:j2 /\*(tʰi2 i1)  
 then 2sg invite  
 ‘then invite (him)!’  
 d. (ve2 ben1) ŋwa:j4 → (ve2 ɲ1) ŋwa:j4 /\*(ve2 e1)  
 return side maternal  
 ‘Going to visit the maternal side.’

Among clitics in (9) [ben1] is not a function word but a noun functioning as a classifier. It might appear that these clitics are independent of the host, i.e., each is a clitic without a host, because they do not behave like those in (8) because they do not surface as lengthened parts of the vowels/glides of the host. However, their surface forms are still determined by their hosts. If the host ends in an obstruent and the clitic has a final nasal, the clitic must give up its original nasal and is assimilated to the final obstruent of the host. The two sentences in (10a) and (10b) have the same meaning except for a different form of self-reference. Since the host in (10a) ends in a glide, the clitic retains its original velar nasal, in this case a plain velar not a labialized one. In (10b), the host ends in a palatal and the clitic surfaces as a palatal, not as a velar.

- (10) a. (toj1 kuŋ<sup>m</sup>6) ɲa:j3 → (toj1 ŋ6) ɲa:j3  
 1sg also intimidated  
 ‘I am also intimidated.’  
 b. (ʔaŋ1 kuŋ<sup>m</sup>6) ɲa:j3 → (ʔaŋ1 ɲ6) ɲa:j3 /\*(ʔaŋ1 ŋ6) ɲa:j3  
 1sg (familiar) also intimidated  
 ‘I am also intimidated.’

The clitic can keep its original nasal only if the nasal is in the coda. If the nasal is in the onset, it often does not survive through cliticization. The clitic just surfaces as a lengthened vowel or glide as in (11).

- (11) a. (ʔa:j1 ma2) ɲɤ3 → (ʔa:j1 j2) ɲɤ3  
           who. Emp. remember  
           ‘Who can remember (that)!’  
       b. tɤ:j3 (dɤw1 nɔ3) tɤ:j3 → tɤ:j3 (dɤw1 w3) tɤ:j3  
           come. where. 3sg.come  
           ‘(I do not care) how it turns out’

Ngô 1984 shows the complementizer [ma2], which has an initial nasal, in *ai mà thương anh* [(a:j1 ma2) tʰiɤŋ1 aɲ1] ‘who could love you?’ with two possibilities after reduction: a nasal with its tone [a:j m2] or just a tone left stray [a:j 2]. He does not transcribe or mention whether the nasal is syllabic, and in the second possibility he offers no comment on how the tone is realized. In the first case, there can be no tone left stray in Vietnamese. Tone 2 must link to the lengthened part of the final glide which is not shown. In the first case [m2], this could be another type of reduction in fast speech for function words (just as vowels in English are often reduced to schwa in fast speech), and it could perfectly be a normal choice here, and an entirely different phenomenon. Note that all three examples of this type given in Ngô share the same environment, i.e. the host on the left ends with a vowel or a glide. The result, however, will be different if this environment is lost: if the host ends in a nasal with a different place, the clitic has no choice, but occurs as a homorganic syllabic nasal, e.g., *ăn một tí* [(an1 mot8) ti3] ‘eat a little’ becomes [(an1 ɲ2) ti3]/\* (an m8), *học một tí* [(hɔkʰ7 mot8) ti3] ‘study a little’ becomes [(hɔkʰ7 ɲᵐ2) ti3]/\* (hɔkʰ7 m8).

### 3.2.3 Hosts with more than one clitic

There are cases in which a host can have more than one clitic. A few examples are given in (12). In (12a) both words [dɔ3] ‘that’ and [tɤ:j3] ‘until’ cliticize leftward onto [hom1] ‘day,’ which ends in a labial nasal. The clitics surface as syllabic [m] with their tones. In (12b) two clitics attach leftward to the host [ʔoŋᵐ1] ‘man’ and surface as a syllabic labialized velar. Note that tone 8 in [mot] cannot occur with a sonorant: tone 2 surfaces instead. Although not common, there are hosts with three clitics, as in (12c). Three words [ʔɤ:j3], [tʰi2], and [la2] attach to the same host, [ciɤ1], which ends in a central diphthong. Note that all three clitics surface as the second element of the diphthong.

- (12) a. ti2 (hom1 dɔ3 tɤ:j3) zɤ2 → ti2 (hom1 ɲ3 ɲ3) zɤ2  
           since day that until now  
           ‘Since that day until now.’

- b. cɔ1 (ʔoŋ<sup>m</sup>1 ʔɤj3 mot8) ta:t7 → cɔ1 (ʔoŋ<sup>m</sup>1 ɲ<sup>m</sup>3 ɲ<sup>m</sup>2) ta:t7  
 give man that one slap  
 ‘Give him a slap!’
- c. buɤj5 (ciɤ1 ʔɤj3 t<sup>h</sup>i2 la2) met4 → buɤj5 (ciɤ1 ɤ3 ɤ2 ɤ2) met8  
 cl noon emp comp be tired  
 ‘It is tiring around noon.’

Note that in (12) if only one clitic is reduced, it would be the one that is closest to the host, i.e. [dɔ3], [ʔɤj3], and [ʔɤj3] in 12a, 12b and 12c, respectively. The fact that when there is more than one clitic, all clitics associate to the same host to their left suggests that, as in K<sup>w</sup>ak<sup>w</sup>ala (Anderson 2005), Vietnamese cannot begin a phonological phrase with a clitic.

In summary, this section has described how clitics surface when all segments are deleted. The right edge of a host determines whether the clitic surfaces as a syllabic nasal or as a lengthened part of a vowel/glide. In the special group in which a clitic has a final nasal and the host ends in a vowel/glide, the clitic surfaces with its nasal. Finally, a host can have more than one clitic attached to it.

The earlier definition of a Vietnamese clitics can now be revised: a clitic is a word that cannot stand on its own. It must associate to a host on its left, regardless of its morphemic status. Only when there is no host to the left does it associate rightward. A clitic is prosodically deficient in that it can occur with a syllabic nasal only when it follows an obstruent.

### 3.2.4 Directionality of cliticization

This section discusses the directionality of cliticization. Clitics attach to the host on their left except when the clitic is phrase-initial when it must attach rightward. If there is a pause before a clitic, the clitic becomes a phrase-initial word and it must attach to the host on its right. This pause is sometimes very brief.

Among clitics the indefinite quantifier [mot8] ‘one/a’ and the adverb [la:m sa:w] ‘how’ are flexible with respect to direction. They can cliticize either leftward or rightward, as shown in the examples in (13). [mot] in (13a) cliticizes rightward, assimilates to the labiodental [f], and surfaces as a labiodental nasal, which is not a phoneme in Vietnamese. In (13b) it can attach to a host on either side.

- (13) a. nɔ3 sɤ4 ca1 (mot8 fɛp7) → nɔ3 sɤ4 ca1 (a1 fɛp7) OR (ɲ2 fɛp7)  
 3sg be afraid father Cl rule  
 ‘He is under his father’s thumb.’
- b. (kɔ3 mot8) ʔoŋ1 → (kɔ3 ɔ2) oŋ1 OR (ɲ2 ʔoŋ1)  
 have. one. man  
 ‘There was a man.’
- c. ba:j2 vɤ5 la:m2 sa:w1 → (vɤ5 ɤ2) OR (ɲ2 sa:w1) OR (ɲ2 sa:w1)  
 study how  
 ‘How is (your) school going?’

Such rightward attachment is rare. In (13a), [mot<sup>8</sup> fɛp<sup>7</sup>] belongs to an idiom *sợ* (somebody) *một phép* ‘be under someone’s rule.’ This may be the reason that the indefinite quantifier [mot] ‘one/a’ has lost its freedom to move. In (13b) [mot] cliticizes rightward to the glottal stop but surfaces as a coronal nasal; since the glottal stop does not have a place, the clitic either keeps the place of its final coronal stop or a coronal is inserted as default. In (13c) [la:m] ‘to do’ has three possible surface forms: either a lengthened vowel if it attaches leftward, or a coronal nasal if it attaches rightward, or retention of its original final labial nasal. Although it ends in a nasal, it is not restricted only to the right side, as are the clitics discussed above. Perhaps its being the first element of the whole adverb *làm sao* [la:m sa:w] ‘how’ also has something to do with this. It should be noted that all the above clitics occur after open syllables so there is a choice in directions of cliticization. When both directions are acceptable, the leftward choice still produces a more desirable output. However, if the host ends in an obstruent, this choice disappears and the clitic must surface as a syllabic nasal.

To summarize, with few exceptions clitics attach to hosts on their left. After being cliticized to a host, all segments of a clitic are deleted, except the tone. This tone is realized on a homorganic nasal if the host ends in an obstruent and as the lengthened part of a vowel or a glide if the host ends in a vowel or a glide. In the latter case, if the clitic has a final nasal that nasal is retained.

The following section will present evidence from spectrographic representations to show that clitics are not ‘full’ forms, that a vowel in a clitic is not just a vowel but is a lengthened part of its vocalic host, and that clitics are not always shorter or weaker than their full form counterparts.

### 3.3 Spectrographic representations of clitics

This section will show when a clitic is reduced to a syllabic nasal, there are no vowel formants in the clitic. It will also show that though a clitic may be weaker it is not necessarily shorter than its corresponding full form. Clitics can also occur in slow speech, as well as fast speech, specifically in casual, colloquial speech.

Figure 1a shows the waveform (first row) and the spectrogram (second row) of full forms in the phrase [(viɤt<sup>7</sup> tɤm<sup>2</sup>) bɤj<sup>4</sup>] ‘to write nonsense.’ Figure 1b shows the same phrase after the second syllable cliticizes leftward and surfaces as a syllabic coronal nasal. A comparison of the second rows in the two figures shows that the vowel formants of [tɤm] are quite clear in Figure 1a but in Figure 1b there is only a syllabic nasal. The durations of the host and the clitic are almost the same in the two figures. The tones of the host and the clitic also have the same pitch curves in the two figures.

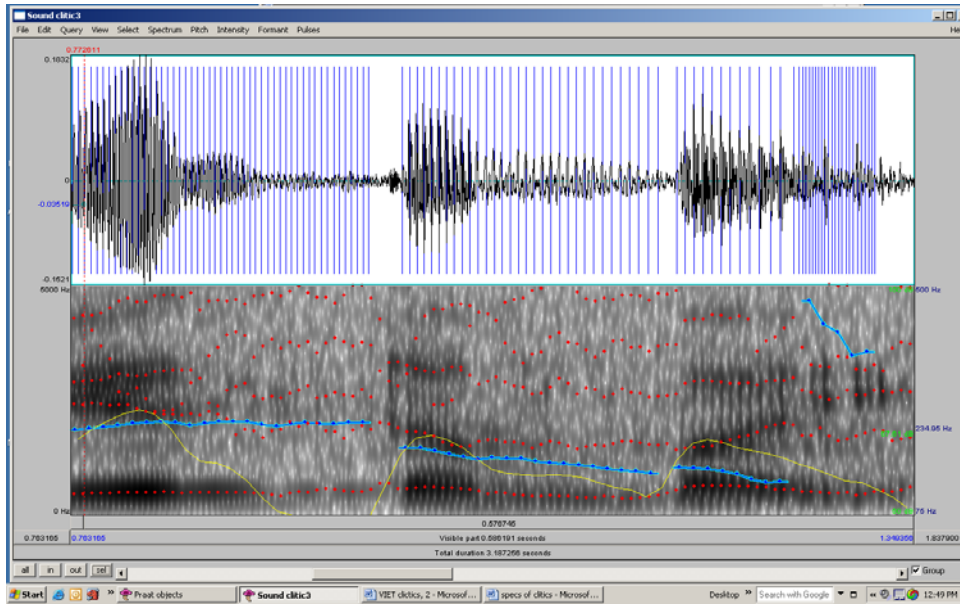


Figure 1a. (vɪɪt7 tɔm2) bɔŋ4

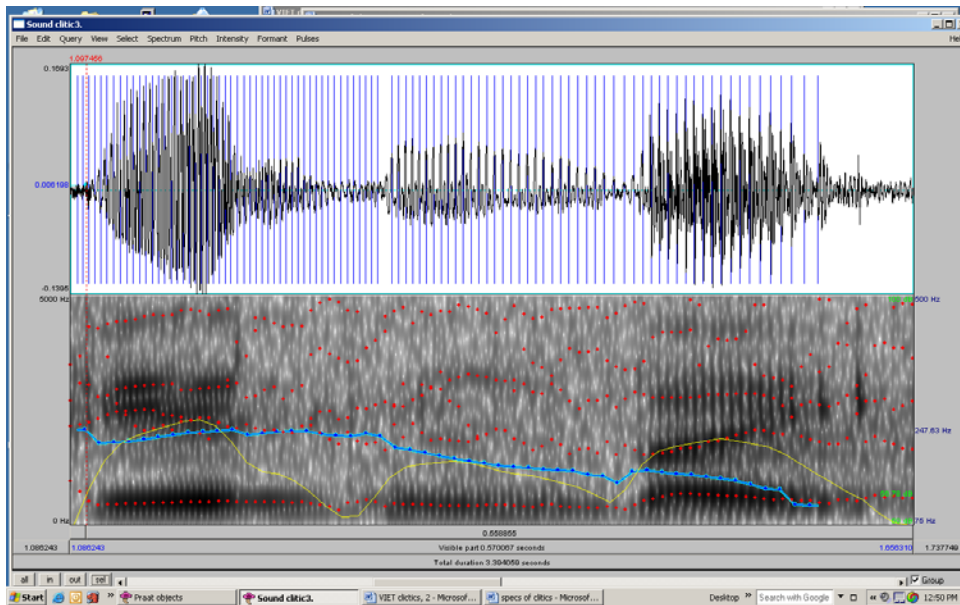


Figure 1b. (vɪɪt7 ɲ2) bɔŋ4

Figure 2a shows the full form and Figure 2b shows the reduced form of a clitic and its host in [mɛw2 ma2] ‘cat that.’ The clitic [ma2] shows vowel formants near the end in Figure 2a but it has no vowel formants when it surfaces as the lengthened glide /w/ in Figure 2b. There is also no clear boundary between the final glide and the clitic, i.e., the lengthened part of the glide, because both the host and the clitic have the same tone. In Figure 2a the host ends where the clitic begins (at the beginning of the shaded area). Figure 2b also shows this boundary. The host and the clitic seem to readjust their lengths, but appear to have the same total length in both figures.



## IS THERE A PROSODIC WORD IN VIETNAMESE?

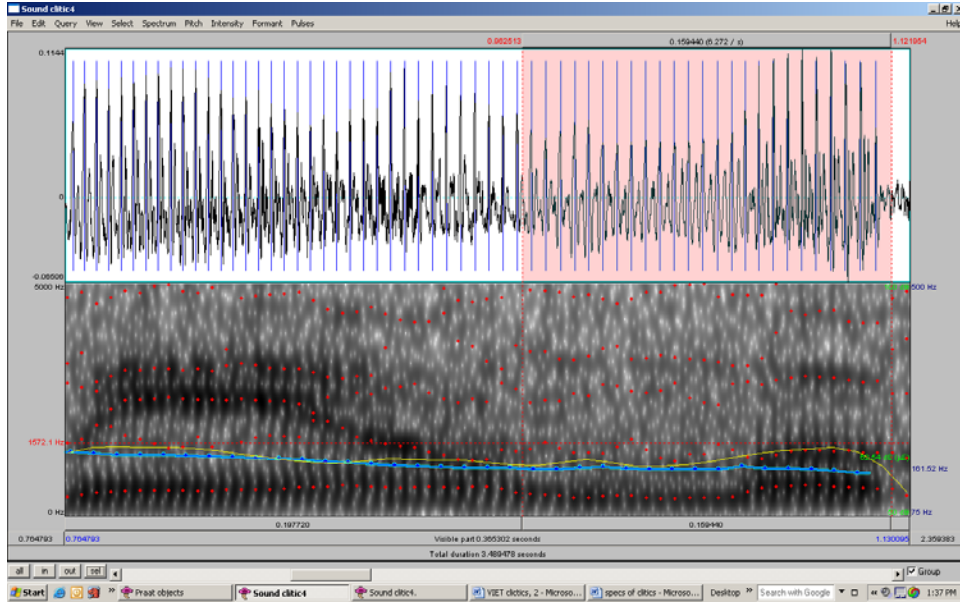


Figure 2a. (mɛ:w2 ma2)

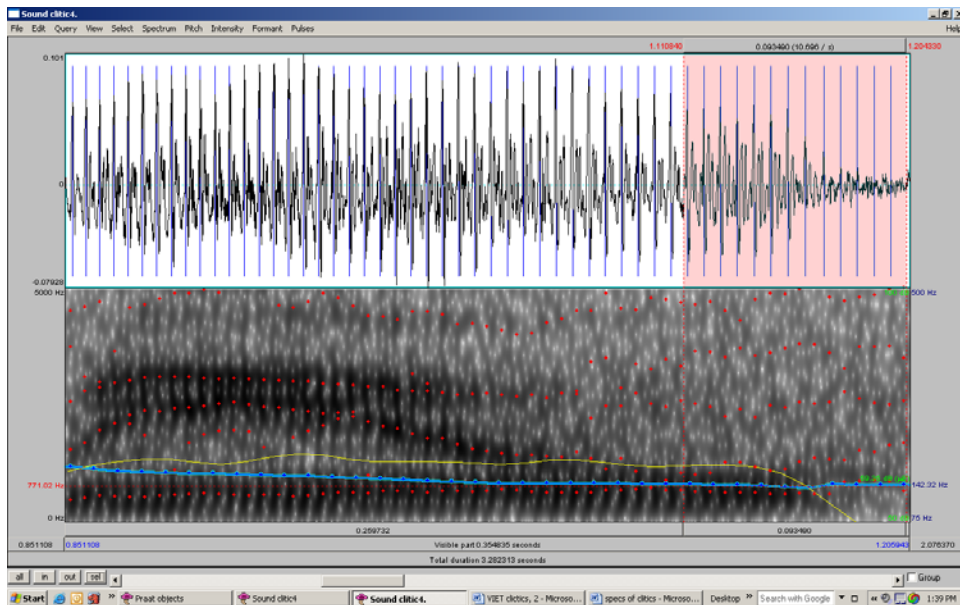


Figure 2b. (mɛ:w2 w2)

Figure 3a shows the full form and Figure 3b shows reduced form of the clitic in the idiom *dù sao đi nữa* ‘anyhow.’ (The first syllable is excluded from the figures.) Figure 3a clearly shows that F1 and F2 of the vowel in [di] are far apart. These formants disappear completely in Figure 3b, where both the host and the clitic are shown in the shaded part, which is also shorter than their counterparts in Figure 3a.

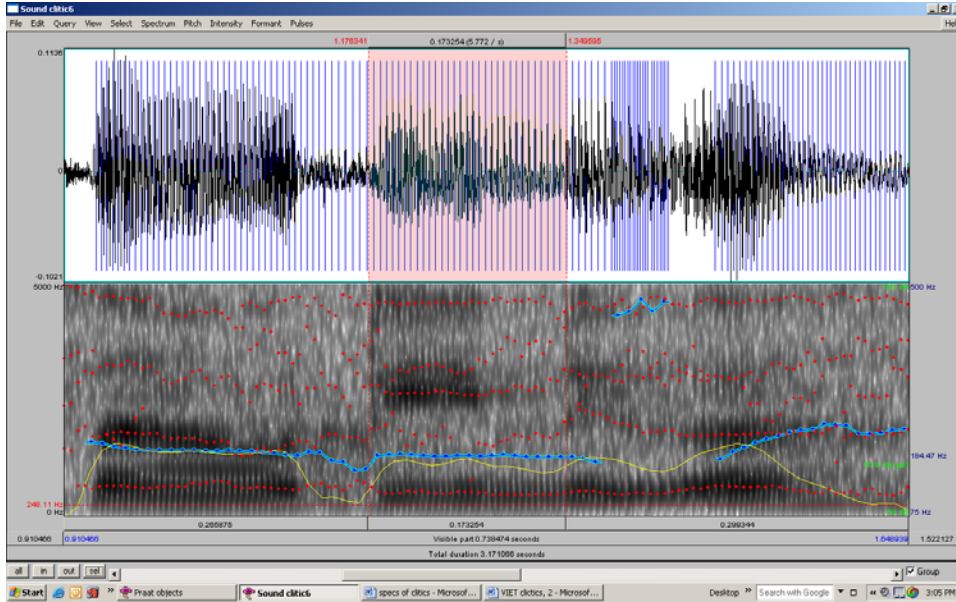


Figure 3a. (sa:w1 d1) nɪʔ6

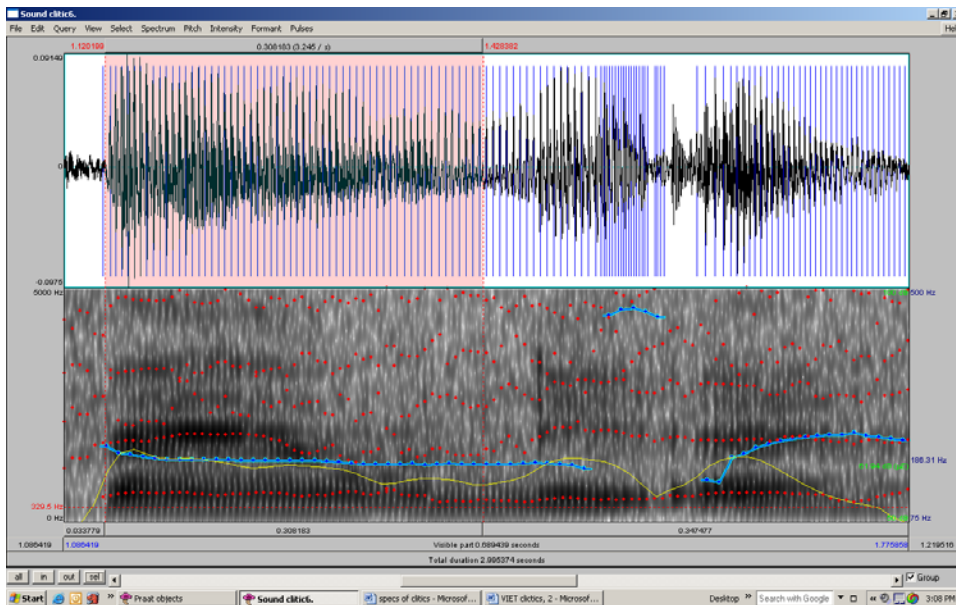


Figure 3b. (sa:w1 w1) nɪʔ6

Figures 4a and 4b show a host and a clitic consisting of two adjacent stops *băt con* [bat3 kɔn1] ‘to catch + cl.’ In Figure 4a the two stops are twice as longer as the final stop in Figure 4b. Figure 4a shows a formant pattern for /ɔ/ not seen in Figure 4b and the syllabic nasal in Figure 4b is much longer than the whole of the full clitic in Figure 4a before it is reduced. This example shows that a clitic is not always shorter than its full form as generally claimed.

## IS THERE A PROSODIC WORD IN VIETNAMESE?

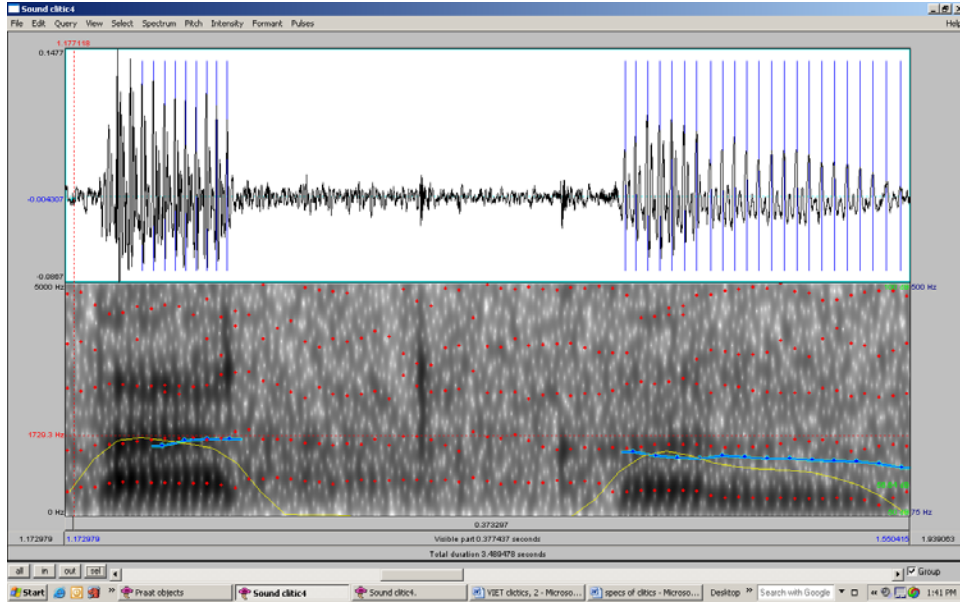


Figure 4a. (bat7 kən1)

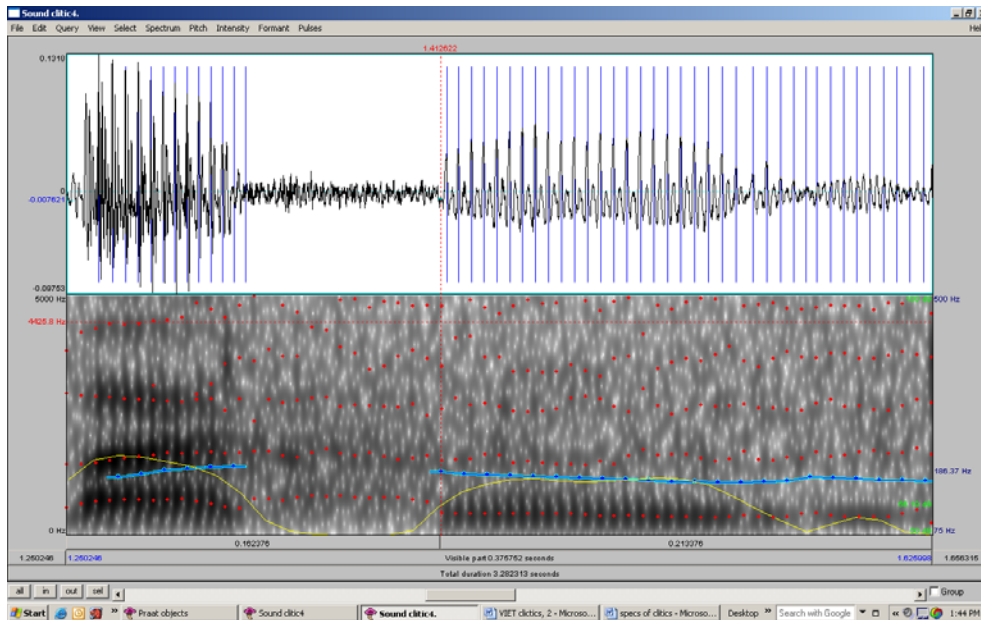


Figure 4b. (bat7 ɲ1)

The next section will describe a phenomenon called ‘contraction’ which is found only in Southern dialects.

### 4. A group of ‘contracted words’ in Southern dialects

In Southern dialects there is a group of words which when combined with a syllable cause that syllable to disappear. This group consists of some kinship terms, locative terms, and time expressions. In this process these words combine with the demonstrative /dɔ/ ‘that,’ which has tone 3. The result is the disappearance of /dɔ/ and the

syllable that survives bears tone 5 regardless of its original tone. For example, [ba2] ‘woman’ + [dɔ3] → [ba5] ‘that woman.’ This phenomenon is mentioned in many studies, e.g., Thompson 1965, Ngô 1984, Hoàng & Hoàng 1975, Bui 1966, Lê 1951.

This process is superficially similar to cliticization but there are important differences. Only one function word participates and it disappears completely, leaving no trace, not even tone. A disyllabic phrase is reduced to a monosyllabic word and there is no place assimilation. The host almost always surfaces with tone 5 regardless of its original tone. This is also not a productive process because only certain words of kinship, location and time participate.

(14) provides some examples of this process. (Note that Southern dialects have three retroflexes /ṭ, ʂ, r/ which do not appear in Northern dialects.) The first syllable is a kinship term (14a-h), a location (14i-m), and a time (14n and o). The second syllable is always the demonstrative *đó* /dɔ3/.<sup>5</sup>

(14)	a. oŋ1 + dɔ3	→	oŋ5	‘that man, he’
	b. ba2 + dɔ3	→	ba5	‘that woman, she’
	c. ko1 + dɔ3	→	ko5	‘that girl, she’
	d. kɤw4 + dɔ3	→	kɤw5	‘that (maternal) uncle/ he’
	e. ku4 + dɔ3	→	ku5	‘that old man/ grandfather/ he’
	f. la:w6 + dɔ3	→	la:w5	‘that old man/ he’
	g. ɲɔ5 + dɔ3	→	ɲɔ5	‘that little (girl)/ she’
	h. a5 + dɔ3	→	a5	‘that girl/woman (inferior)’
	i. ben1 + dɔ3	→	ben5	‘that side/ there’
	j. daŋ2 + dɔ3	→	daŋ5	‘that side/ over there’
	k. ŋwa:j2 + dɔ3	→	ŋwa:j5	‘outside/ out there’
	l. tɕŋ1 + dɔ3	→	tɕŋ5	‘inside/ in there’
	m. tɕen1 + dɔ3	→	tɕen5	‘above/ up there’
	n. hom1 + dɔ3	→	hom5	‘that day’
	o. nam1 + dɔ3	→	nam5	‘that year’

This process is not productive because not any location or time expressions can be contracted this way, e.g. [jɤ2 + dɔ3] ‘that hour’ → \* [jɤ5], [twɤn2 + dɔ3] ‘that week’ → \* [twɤn5]. Note also that the contracted form in (14a) with the meaning ‘that man,’ is equivalent to Northern dialects *ông ấy* [oŋ<sup>m</sup>1 + ɤj3] ‘that man.’ The Northern cliticized equivalent is [oŋ<sup>m</sup>1 ɤ<sup>m</sup>3]. Likewise, [ben1 dɔ3] ‘that side, over there’ appears as [ben1 ɲ3] in the Northern dialect. Unlike the ‘true’ clitics when either a host or a clitic can be part of a disyllabic word, in Southern dialects the words involved in this process are

<sup>5</sup> The Southern demonstrative /dɔ3/ ‘that’ has the same meaning as the Northern /ɤj/, and both have the same tone.

always monosyllabic. (14) also shows that the first syllable can have any tone but the contracted form has tone 5 except in a few rare cases when the first syllable has tone 3.

- (15) a. t<sup>h</sup>im3 + dɔ3 → t<sup>h</sup>im3 /\*t<sup>h</sup>im5 ‘that aunt (paternal uncle’s wife)’  
 b. cu3 + dɔ3 → cu3 /\*cu5 ‘that paternal uncle’

In such cases, we can also see exactly what happens with genuine clitics: the result is a disyllabic word with the tone of the clitic /dɔ/ realized on a vowel or a syllabic nasal, i.e., it would be (t<sup>h</sup>im3 m̩3) in (15a), and (cu3 u3) in (15b) so cliticization is also found in the colloquial speech of Southern speakers. Only the combinations in (14) are special in that each combination contracts to one syllable with tone 5.

In Mixtec, an Otomanguen language, there is a group of clitics in which certain segments and initial syllables are deleted under appropriate conditions (Macaulay 1987), and as a consequence, disyllabic words are reduced to monosyllabic words in fast speech. This group of contracted forms in Southern Vietnamese seems to be similar to that group of clitics in Mixtec. However, only very specific words can participate in this process in this variety of Vietnamese.

Tone 5 in these contracted forms appears to be lexically specified, i.e., it must be listed as an alternative form of the full form [dɔ3], just like reduced auxiliaries such as ‘s (from *is*), ‘d (from *would*), ‘ve (from *have*) in English (Kaisse 1985). In other words, [dɔ3] is deleted, and the tone of the first syllable is replaced with tone 5. Note that all the disyllabic forms on the left column in (14) can occur as an alternative, e.g., [hom1 + dɔ3] → [hom1 dɔ3] ‘that day,’ depending on the style of speech. To a native of this Southern variety of Vietnamese, tone 5 serves as a trace of the deleted demonstrative [dɔ3] and adds its meaning. Consequently, this appears to be an exception to the claim in Ngo 1984, ‘the minimal meaningful unit of grammatical analysis is never smaller than a syllable.’

Ngô 1984 provides another analysis. He explains that contraction does not occur if the noun bears either tone 3 or 5 because these tones would lose their contour distinctiveness when combined with tone 3 of the demonstrative /ɤj3/. Contraction, therefore, is ‘simply a phonetic adjunction of the pitch contour’ (1984: 107). However, Southerners do not use /ɤj3/, as stated in Ngô, but use /dɔ3/ instead, although both words have the same tone. This becomes obvious in non-contracted forms, when the demonstrative is not reduced. Moreover, any explanation based on phonetic contours is inaccurate: tone 4 in Southern dialects has the falling-rising contour, just like tone 5 in Northern dialects (Hoàng 1989, Vũ 1982), but it also undergoes contraction. Tone 6 (neutralized to tone 5 in Southern dialects) can also participate in contraction, as seen in (14f). Syllables with tone 6 have a very low frequency in the language, but (14f) shows that they can too undergo contraction.

This section has presented a phenomenon that is similar to cliticization in Northern Vietnamese dialects; a process in which the demonstrative /dɔ3/ is deleted when it follows a restricted set of lexical items. The contraction in Southern dialects is not cliticization since it is unique to a particular morpheme and the whole syllable is erased including its tone.

## 5. Conclusion

This paper has examined the phenomenon of cliticization in Vietnamese casual speech. Clitics are function words that associate with a host on their left regardless of the morphosyntactic status of that host. The segments of the clitic are lost but the tone remains to be realized on a syllabic nasal or as a lengthened vowel/glide depending on the rhymal structure of the host and of the clitic itself.

If the analysis proposed in this paper is correct, the phenomenon of cliticization provides us with a better understanding of phonology at the level of the foot, i.e., what exactly can constitute a foot. It also raises the issue of whether a prosodic word indeed exists in Vietnamese, a language whose prosodic structure is still very much under-investigated.

## References

- Anderson, Stephen R. 2005. *Aspects of the theory of clitics*. New York: Oxford University Press.
- Benedict, Ruth. 1948. Tonal system in Southeast Asia, *JAOS* 68: 184 - 91.
- Bùi, Đức Tịnh. 1966. *Văn phạm Việt Nam (Vietnamese grammar)*. Saigon: Xuân Thu.
- Cao, Xuân Hạo. 1978. *Trọng âm và các quan hệ ngữ pháp trong tiếng Việt (Stress and syntactic relations in Vietnamese)*. In *Thông báo ngữ âm học*. Ho Chi Minh city: Social Science Institute.
- Cao, Xuân Hạo. 1998. *Tiếng Việt- mấy vấn đề ngữ âm, ngữ pháp, ngữ nghĩa (Vietnamese: some issues in phonology, syntax, semantics)*. Hồ Chí Minh City: Giáo Dục Press.
- Emeneau, M.B. 1951. *Studies in Vietnamese (Annamese) grammar*. Berkeley: University of California Press.
- Hildebrandt, Kristine, Rene Schiering & Balthasar Bickel. 2006. Cross-linguistic challenges for the Prosodic Hierarchy: evidence from Word Domains. Paper given at the 14<sup>th</sup> Manchester Phonology Meeting, May 25-28. University of Manchester.
- Hồ, Lê. 1976. *Vấn đề cấu tạo từ của tiếng Việt hiện đại (On the issue of word formation in modern Vietnamese)*. Hà Nội: Khoa Học Xã Hội.
- Hoàng, Cao Trọng. 1989. Thanh điệu Việt qua Giọng Địa Phương Trên Cú Liệu F0 (The F0 values of Vietnamese tones in dialects). *Ngôn Ngữ (Linguistics)*, 4: 1-17.
- Hoàng, Tuệ and Hoàng Minh. 1975. Remarques sur la structure phonologique du Vietnamien. Hà Nội: *Vietnamese Studies*, 40. 67-97.
- Honey, P.J. 1989. Vietnamese speakers. In Michael Swan and Bernard Smith (eds). *Learning English – A teacher's guide to interference and other problems*. Cambridge: Cambridge University Press, 238-51.
- Kaisse, E. 1985. *Connected speech: the interaction of Syntax and Phonology*. New York: Academic Press.
- Liberman, Mark Y. & Alan S. Prince. 1977. On stress and linguistic rhythm. *Linguistic Inquiry* 8. 249-336.
- Macaulay, Monica. 1987. Cliticization and the morphosyntax of Mixtec. *International Journal of American Linguistics* 53. 119-35.
- Ngô, Như Bình. 1999. Khắc phục một số trở ngại trong việc dạy tiếng Việt cho người bản ngữ tiếng Anh ở giai đoạn đầu (How to help English speakers to overcome obstacles in learning Vietnamese at the beginning level). In *Ngôn Ngữ (Linguistics)*, No 3, Hà Nội, 73-78.
- Ngô, Thanh Nhân. 1984. *The syllabeme and patterns of word formation in Vietnamese*. PhD Dissertation. New York University.
- Nguyễn, Công Đức. 1994. Về kết cấu song tiết lấy âm tiếng Việt hiện đại (Structure of disyllabic reduplication in modern Vietnamese). *Ngôn Ngữ (Linguistics)*, No 4. Hà Nội. 47-55.
- Nguyễn, Đình-Hoà. 1990. Vietnamese. In Bernard Comrie (ed.). *The world's major languages*. London: Oxford University Press. 777 - 96.
- Phạm, Andrea H. 2003. *Vietnamese tone: a reanalysis*. New York: Routledge.
- Phạm, Andrea H. Vietnamese as an iambic language. Ms, University of Toronto.
- Peperkamp, Sharon. 1997. *Prosodic Words*. Den Haag: Holland Academic Graphics.
- Selkirk, Elizabeth O. 1978. On prosodic structure and its relation to syntactic structure, in T. Freheim, ed. *Nordic Prosody II*. Trondheim: Tapir, 111-140.

- Selkirk, Elizabeth O. 1996. The prosodic structure of function words. In Jill N. Beckman, Laura Walsh Dickey & Susanne Urbanczyk (eds). *Papers in Optimality Theory*. Amherst: UMASS, 439-470.
- Selkirk, Elizabeth & Tong Shen. 1990. Prosodic domains in Shanghai Chinese. In Sharon Inkelas & Draga Zec (eds). *'The phonology-Syntax Connection.'* Chicago: University of Chicago Press.
- Thompson, Laurence. 1965. *A Vietnamese grammar*. Seattle: University of Washington Press.
- Vũ, Thanh Phuong 1982. Phonetic properties of Vietnamese tones across dialects. Canberra, *Papers in South East Asian Linguistics*, 55-76.
- van der Leeuw. 1997. Clitics – Prosodic Studies. Den Haag: Holland Academic Graphics.
- Wackergael, Jacob. 1892. Über ein Gesetz der Indogermanischen Wortstellung. *Indogermanische Forschungen* 1. 333-436.
- Zwicky, Arnold M. 1985. Clitics and particles. *Language* 61. 283-305.