



琉球大学学術リポジトリ

University of the Ryukyus Repository

Title	Clitics in Irabu Ryukyuan
Author(s)	Shimoji, Michinori
Citation	International journal of Okinawan studies, 4(1): 51-79
Issue Date	2013-09-02
URL	http://hdl.handle.net/20.500.12000/35886
Rights	

Clitics in Irabu Ryukyuan

Michinori Shimoji*

1. Introduction

The aim of this study is to describe clitics in Irabu Ryukyuan, a Southern Ryukyuan variety of the Japonic Family spoken on Irabu Island, Okinawa. Based on the widely recognised assumption that clitics are phonologically dependent word-like elements that are distinct from words and affixes, I identify the following elements as clitics in Irabu: case markers, quasi-quantifiers such as = *mai* ‘too’ and = *tjaaki* ‘only’, information-structure markers such as topic markers and focus markers, modal markers and discourse markers.

Morphosyntactically speaking, Irabu clitics are of the type often referred to as “phrasal affixes” (Zwicky 1994, Anderson 1992, and many others) or “categorically unrestricted bound formatives” (Bickel and Nichols 2007). They mark grammatical features of phrases (or heads of phrases) such as case, information-structure status (topic or focus), mood, etc. As such, a careful argumentation is required for the distinction between clitics and inflectional affixes, the latter of which also designate grammatical features. I argue that clitics are distinct from inflectional affixes in their ability to be in construction with phrases or clauses (rather than stems), i.e., their distributions at a syntactic level just as in the case of words. The word-like character of clitics is thus reflected in the surface fact that they exhibit a lower degree of selectivity than do affixes (Carstairs 1981, Zwicky and Pullum 1983, Aikhenvald 2002, Bickel and Nichols 2007, Haspelmath 2011, and many others). On the other hand, clitics are distinct from words as well, since they are placed by very simple distributional principles and since they are partially immune to syntactic rules, especially the movement rule.

Phonologically speaking, Irabu clitics comprise ones that constitute a single phonological word with the host (Internal clitics), and ones that do not (External clitics). This means that there needs to be a distinction between the phonological word domain and the domain consisting of a phonological word and an External clitic (as opposed to two phonological words that constitute a phonological phrase). The latter domain corresponds to the “Clitic Group” as postulated by Selkirk’s (1984) and Nespor and Vogel’s (1986) Prosodic Hierarchy, in which the Clitic Group is situated between the Phonological Word and the Phonological Phrase.

This paper is organised as follows. In Section 2, I discuss the notion of clitic as under-

* Associate Professor, Kyushu University 九州大学准教授

stood in the typological and theoretical literature and define the clitic in Irabu. Sections 3 and 4 examine the morphosyntactic properties of Irabu clitics. Sections 5 and 6 turn to the phonological aspects of Irabu clitics. Section 7 draws a conclusion.

2. Identification of clitics

A fairly common view on the notion of clitic in linguistic typology is that a clitic is a phonologically dependent but syntactically word-like element (Matthews 1974, Anderson 1985, Haspelmath 2002, Aikhenvald 2002, Booij 2005, Dixon 2009), a definition which stems from the older Indo-Europeanists' tradition (and which is also adopted by the *Oxford English Dictionary*; see Anderson 2005: 1).

The agreement as to the definitional character of clitics in the cross-linguistic sense centres on their phonological deficiency, and their grammatical peculiarities that make them word-like but still distinguished from full-fledged words are more controversial. After a thorough review on clitics in the typological and theoretical literature, Aikhenvald (2002: 43) clearly states that the “consensus appears to be that clitics are morphemes which are prosodically deficient or unusual in certain ways.” However, she does not give as clear a generalisation about the grammatical characteristics of clitics, giving a wide range of parameters according to which the grammatical properties of clitics may vary cross-linguistically. Anderson (2005: 1) states that “a clitic is something that is not integrated into the sentence in the way ‘normal’ words are, and/or not integrated into words in the way affixes are.” This negative characterisation of clitics against words on the one hand and affixes on the other makes a positive identification of an element as a clitic especially difficult in descriptions of individual languages and in researches in theoretical orientation (Zwicky 1977, 1985, 1994, Zwicky and Pullum 1983, Anderson 1992, 2005). Very often, various kinds of dependent elements which have some sort of word-like character and affix-like character are collectively called “clitics,” making this notion less meaningful in linguistics (Zwicky 1995).

However, it is often pointed out that clitics can be a significant unit that can be distinguished from words and from affixes by the existence of their “special syntax” (Spencer and Luis 2012: 44). For example, the so-called “second position clitics” are put after the first constituent of a sentence (‘Wackernagel’ position). Other languages put clitics on other specific syntactic positions of a sentence (e.g. sentence-initial, as in Kwak’wala determiners, Anderson 1992). Such simple distributional principles make the fixed-position clitics not word-like. The syntactically special character of clitics is often reflected in their inability to be subject to certain syntactic rules that ordinary word would obey, such as movement, replacement and ellipsis (Zwicky and Pullum 1983, Zwicky 1985, Haspelmath 2011), which indicates that clitics occur on a special layer of syntactic structure. For example, the French definite article is like a full-fledged word in this language, since it may be subject to the coordination rule, which conjoins coordinants by *ou* ‘or’, as in *On peut dire le ou la pamplemousse* (‘One can say *le* or *la* pamplemousse (grapefruit)’), Miller

1992: 151). However, it is not like a word since it is immune to the ellipsis rule, i.e. they cannot be ellipted in coordinate constructions and must be repeated like *les garcon de Paris et les fille de Milan* (cf. **les garcon de Paris et filles de Milan*; Miller 1992: 12).

The strongest hypothesis about the grammatically special status of clitics is given in Zwicky and Pullum's (1983) and Zwicky's (1985) model, according to which the clitic is claimed to occur after all syntactic rules apply, with the theoretical prediction that no syntactic rule exists which targets the construct word + clitic. In this model, therefore, clitic attachment is expected to occur at a surface or a near-surface level.

The position taken in this paper in identifying clitics is not so strict as Zwicky's model, singling out those phonologically dependent elements which have special syntactic properties that cannot be appropriately described by referring them simply as words or affixes. As an initial approximation, Table 1 summarises the distinction between words, clitics and affixes in Irabu. As will be extensively discussed in Sections 3 and 4, Irabu clitics are like words in that they occur at a syntactic level, combining with phrases or clauses rather than stems, which clearly distinguish clitics from affixes, which are by definition morphological in nature. On the other hand, their distributions are regulated by fairly simple syntactic principles that are not found in ordinary words. Their syntactically special status is also evident in the fact that they are partially immune to certain syntactic rules that target words such as movement, indicating that they occupy a special position (or they occur at a special layer) within the syntactic structure. Furthermore, I will show that a certain set of discourse markers clearly occur after all syntactic rules apply and clearly occur at a near-surface level (in terms of the phonological criteria that I will examine in Sections 5 and 6), thus perfectly following the prediction that Zwicky's model makes about their notion of clitic as a theoretical construct.

Table 1: Word, clitic and affix in Irabu

		Word	Clitic	Affix
Morphosyntax	Distributional principle	Syntactic (complex)	Syntactic (simple)	Morphological
	Syntactic rule	Applicable	Partially applicable	Non-applicable
Phonology		Independent	Dependent	Dependent

The above characterisation of clitics excludes an element from the candidate list of clitics if it is merely phonologically dependent and has an ordinary syntactic characteristic as a word, or if it is phonologically independent even if it has special syntactic properties like simple distribution and partial immunity to syntactic rules. Thus, the following elements are classified as words rather than clitics:

- (1) Phonologically dependent words
 - a. Interjections *mmja* 'well' and *hira* 'hey; you know', which do not constitute

independent phonological words on their own, leaning on their preceding or following elements. However, unlike clitics, their distributions cannot be easily described by simple principles and are like other interjections that are phonologically independent such as *naugara* ‘well’ and *gammja* ‘oh my god!’

- b. The first person singular pronoun *ban* is realised as a bound form *ba* in its nominative form, as in *ba=ga* (1SG=NOM). This form is never treated as an independent phonological word on its own. However, there is no special syntactic pattern exhibited by this form, behaving exactly like other free pronouns.
- c. The light verb *as* ‘do’ (and its inflected forms) has a bound form *s*, which is never treated as an independent phonological word on its own, as in *nau s-tar?* ‘what did (you) do?’ (what do-PST).¹⁾ However, its distributional pattern is describable by assuming that it is simply a reduced variant of the full-fledged form. No special syntax is observed.
- d. The formal nouns *kja* ‘when’, *njaa* ‘like’ and *jau* ‘in order to’ (see Shimoji 2008 for their detailed description), which all occur as heads of noun phrases that carry adnominal clauses. They are all bound and phonologically dependent in the sense that they do not constitute a phonological word on their own. Syntactically, however, they are very much like ordinary nouns that carry an adnominal clause, since they require the predicate verb of the adnominal clause to be inflected as an adnominal form, and they, as heads of noun phrases, carry case markers as ordinary nouns do.

(2) Phonologically independent words with special syntactic properties

- a. The copula verbs *ar* (and its inflected forms) and *jar* have a special syntactic pattern in terms of its positioning (it occurs at the final position of a predicate nominal phrase) and its inability to be subject to the movement rule. However, they are phonologically independent and so do not qualify as clitics.
- b. The auxiliary verbs like *ur* (progressive), *ar* (resultative), *njaan* (perfect), etc. (and their inflected forms) have special syntactic patterns, as they occur at the final position of a verbal predicate phrase and they cannot be moved independently. However, they are phonologically independent and so do not qualify as clitics.

It is likely that these elements, especially those in (2), are on their way to becoming clitics diachronically.

Conversely, the following elements are treated as affixes rather than clitics, as their distribution and other properties are best described by referring to inflectional morphology.

(3) Post-inflectional affixes

The formal-noun-like *-su(u)* ‘person; thing; fact’ and the forms that historically

developed from it (such as the concessive subordinate marker *-suga* ‘though’ and sentence-final discourse marker *-suuda* ‘isn’t it?’) are clitic-like in that they are phonologically dependent and that their distribution appears to be syntactically determined: they occur at clause-final position. The following examples illustrate how *-su* (a) and *-suga* (b-c) are used.

- a. *munu=u s-sa-n-su(u)=nkai ck-as-i.*
 thing=ACC know-IRR-NEG-person=ALL listen-CAUS-IMP
 ‘Inform (it) of those who do not know things.’
- b. *ba=a mii-t-tar-ruga=du ffa-gama-ta=nu mii-tar=ca.*
 1SG=TOP look-NEG-PST-though=FOC child-DIM-PL=NOM look-PST=HS
 ‘I didn’t look (at the dance), but (my) kids looked (at the dance), they said.’
- c. *aparagi-pztu jar-ruga=du=i, mmja zin=na njaan=ti.*
 beautiful-person COP-though=FOC=CNF well money=TOP not.exist=QT
 ‘(She) is a beautiful person, you know, but (she) does not have money.’

However, they do not simply attach to a clause but always require the clause to end in a verb or an adjectival verb. Note that in (c) above, the concessive *-suga* requires the clause to end in the copula verb even when the copula is unnecessary when a clause is in non-past tense. This is because the source structure of these forms is the adnominal clause structure in which the head noun is *su(u)*, as in (a). As a result, their phonological hosts are always verbal (verb or verbal adjective). This category-specific property of these forms is typical of morphological elements, i.e. affixes, the distribution of which is best explained by referring to inflectional morphology. They attach to a fully-inflected word form as extra elements.²⁾

As will be discussed in the following sections, Irabu clitics comprise case markers, quasi-quantifiers, information-structure markers, modal markers and discourse markers.

(4) Irabu clitics

- a. Case markers: *=ga* (nominative/genitive 1), *=nu* (nominative/genitive 2), *=ju* (accusative 1), *=ja* (accusative 2), *=n* (dative), *=nkai* (allative), *=sii* (instrumental), *=kara* (ablative), *=gami* (limitative ‘until; as far as’), *=tu* (comitative), *=jarruu* (comparative)
- b. Quasi-quantifiers: *=mai* (even; too), *=tjaaki* (only), *=bakaar* (only), *=kara* (primarily), *=gami* (exactly)
- c. Information-structure markers: focus markers *=du* (declarative focus), *=ru* (Yes-No interrogative focus), *=ga* (WH interrogative focus), topic markers *=ba* (object topic) and *=a* (non-object topic)
- d. Modal markers: the dubitative *=bjaam* ‘I wonder (if)’, another dubitative marker *=gagara* ‘I wonder (what, how, etc.)’, uncertainty *=paz* ‘maybe’, assertive *=dara*, emphatic *=doo* and reserved emphasis *=saa*.
- e. Discourse markers: *=ju* (corrective; *=Y* ‘(not X) but Y’), *=ca* (hearsay), *=da*

(presentational; ‘how about...?’), =*ti(i)* (quotative), =*e(e)* (question) and =*i(i)* (confirmative, which functions like English tags).

3. Morphosyntactic aspects of Irabu clitics (1): Distributional principles

3.1 Clitics and inflectional affixes

Clitics in Irabu are in construction with phrases or clauses. Their distributions are describable with much simpler principles than words, e.g. “a case clitic occurs at the left-most part of a noun phrase.” Zwicky (1985) notes this simple distribution as an indication of a clitic (and affix) as opposed to a word. That is, the simple distribution exhibited by clitics is also characteristic of affixes. Therefore it is necessary to carefully distinguish clitics from affixes, especially *inflectional* affixes. We actually find a number of characteristics that are shared by clitics and inflectional affixes. One such commonality is morphological regularity: the simple distributional principle of inflectional affixes necessarily allows them to occur with almost any stem of the specific class for which it is affixed. For example, verbal inflectional affixes occur with all lexical verbal stems. This regularity (or a lack of “arbitrary gaps” in Zwicky and Pullum’s 1983 terms) is also typical of clitics, which regularly combines with a specific type of phrase without respect to the semantic property of the head word.

However, the difference between a clitic and an (inflectional) affix becomes obvious when attention is paid to the *phonological* hosts to which clitics and affixes are attached. Table 2 classifies Irabu clitics according to the type of syntactic hosts and phonological hosts. The syntactic host of a clitic may be one or more of the following kinds of phrases: noun phrase, argument phrase (which may maximally consist of a noun phrase plus a case marker, quasi-quantifier and information-structure marker), adjunct phrase (which may be a single adverbial word or an adverbial clause), predicate phrase (either verbal, adjectival or nominal) or an utterance (which includes all the above kinds of phrases and other fragments such as conjunction, interjection and bits of phrases). The phonological host of

Table 2: Clitics in terms of syntactic and phonological hosts

	Syntactic host	Phonological host							
		Nominal	Verbal	Numeral	Adjective	Adverb	Conjunction	Adnominal	Interjection
Case markers	Noun phrase	*	*	*	*				
Quasi-quantifiers	Argument, adjunct or predicate phrase	*	*	*	*	*			
Modal markers	Predicate phrase	*	*	*	*	*			
I-S markers	Any type of phrase	*	*	*	*	*	*		
Discourse markers	Any type of utterance	*	*	*	*	*	*	*	*

a clitic may be one or more of the following parts of speech: Nominals (nouns, pronouns, interrogatives, adjectival nouns, etc.), Verbals (verbs and adjectival verbs), Adjective, Numeral, Adverb, Conjunction, Adnominal and Interjection.³⁾ As is clear from the table, since a clitic attaches phonologically to the last element of a phrase, the phonological host is a chance neighbour rather than a specific type of stem as in the case of affixes. For example, the syntactic host of case markers is a noun phrase, which may end in a nominal, verbal, adjective, and numeral, so the phonological hosts of case markers vary according to whether a noun phrase ends in one of these parts of speech.

3.2 Case markers

A case marker is attached syntactically to a noun phrase. It is quite clear that a case marker is in construction with a phrase (syntactically) rather than a noun (morphologically), given that a case marker scopes over an entire noun phrase and occurs per phrase. For example, in (5) the nominative case marker =*ga* is attached to an entire phrase (indicated by the bracket) that consists of an adnominal modifier, a head noun and a numeral in apposition with the head noun.

- (5) [*unu sinsii taukjaa*]=*ga=du* *if-kutu*.
 that teacher one.person=NOM=FOC go-OBL
 ‘That teacher alone should go.’

Unlike languages such as Norwegian where each coordinand of a noun phrase consisting of coordinate nouns is case-marked (e.g. *han og meg* (he.NOM and I.ACC) ‘he and I’; Johannessen 1998:1), Irabu puts a case marker per phrase, indicating that the case marking as well as coordination is syntactic in nature. Thus, the first coordinant *vva* is not marked by the nominative case, but marked instead by the comitative marker =*tu*.

- (6) [*vva=tu ban*]=*ga=du* *if-kutu*.
 2SG=COM 1SG=NOM=FOC go-OBL
 ‘You and I should go.’

Construction with phrases justifies the syntactic (word-like) nature of case markers in Irabu (see also Klavans 1985 for the argument in favour of such an analysis in theoretical research). There is further evidence for the nominative case being in construction with a phrase. The phonological host of the nominative case marker is the first person pronoun *ban*. When this pronoun alone constitutes a noun phrase, as in (7), idiosyncratic allomorphy is induced and we get *ba=ga* rather than **ban=ga*.

- (7) *ba=ga=du* *if-kutu*.
 1SG=NOM=FOC go-OBL
 ‘I should go.’

The fact that the personal pronoun changes its form depending on whether it forms a simplex noun phrase (7) or a complex one (6) indicates that a case marker is attached *after* a noun phrase is formed, not *before* a noun phrase is formed irrespective of the syntactic context in which the pronoun is put. The same kind of allomorphy is found in accusative marking. The first person pronoun is realised as *banu-* when the accusative marker =*ju* is attached (and we get *banu=u*). However, if the noun phrase is a complex phrase consisting of two conjoined nouns as in (8), the pronoun is the underlying *ban*.

- (8) [*vva=tu ban*]=*nu=du saar-i-i if=paz*.
 2SG=COM 1SG=ACC=FOC take-INF-SEQ go=maybe
 ‘(They) will ctake you and me.’

The phonological host of a case marker varies according to the element which comes phrase-finally. Since a noun phrase in Irabu mostly ends in a head noun, the phonological host of a case marker is usually a noun, as in (9), where the accusative case marker =*ju* is syntactically attached to the bracketed noun phrase and is phonologically attached to the noun *kutu* ‘fact’.

- (9) [*kai=ga ssagi=u as-tar kutu*]=*u=du cf-tar*.
 3SG=NOM bridal=ACC do-PST fact=ACC=FOC hear-PST
 ‘(I) heard the fact that he did a bridal.’ [phonological host: noun]

However, the phonological host may also be a verb or adjectival verb if the noun phrase contains a headless relative clause, or a numeral if the noun phrase consists of the head noun and an appositional numeral.

- (10) [*kai=ga ssagi=u as-tar*]=*ru=du cf-tar*.
 3SG=NOM bridal=ACC do-PST=ACC=FOC hear-PST
 ‘(I) heard (the news that) he did a bridal.’ [phonological host: verb]
- (11) [*imi-kar*]=*ru=kara muc-i-kuu*.
 small-ADJ.NPST=ACC=first carry-INF-come.IMP
 ‘Bring (those which) are small first.’ [phonological host: adjectival verb]
- (12) [*ui pitic*]=*cu=mai as-irai-n*.
 3SG one=ACC=even do-POT-NEG.NPST
 ‘(He) cannot even do that.’ [phonological host: numeral]

3.3 Quasi-quantifiers

Quasi-quantifiers (or “restrictives” in Martin’s 1975 terms for similar morphemes in Japanese) comprise =*tjaaki* ‘only’, =*bakaar* ‘only’, =*mai* ‘even; too’, =*gami* ‘even’ and =*kara* ‘primarily’. They attach syntactically to argument phrases or predicate phrases, but

in restricted cases to adjunct phrases as well. Whereas case markers attach simply to the left-edge of noun phrases, the syntactic positioning of quasi-quantifiers is slightly more complex.

First, quasi-quantifiers may attach syntactically to argument phrases, adjunct phrases and predicate phrases. Their distributional principles are still argued to be much simpler than words, as they usually attach to the left-edge of these kinds of phrases. However, if the syntactic host is a predicate phrase, they must occur within a complex predicate phrase (which consists of a lexical component and an auxiliary component) and after the lexical component (predicate nominal/adjective or lexical verb). For example, in (13a) the quasi-quantifier *=mai* ‘too; even’ is syntactically attached to the entire predicate phrase (in brackets) that consists of a lexical verb *snii* ‘die’ (the sequential converbal form) and the perfect auxiliary *uf*, but it is phonologically attached to the lexical verb (underlined) rather than to the predicate-final auxiliary. In (13b, c) the predicate phrase to which the quasi-quantifier *=gami* (EMP) is syntactically attached is a nominal predicate phrase (bracketed), in which the lexical component is filled by either a nominal or adjective. The clitic is therefore attached phonologically to a nominal or an adjective (underlined).

- (13) a. *nnama=gami=a [sn-i-i=mai=du uf]=i.*
 now=EMP=TOP die-INF-SEQ=too=FOC PERF=CNF
 ‘By now, (we would) have died, eh?’ [phonological host: verb]
- b. *kuri=a [gakusja=gami=du jar].*
 3SG=TOP scholar=EMP=FOC COP
 ‘He is a scholar.’ [phonological host: noun]
- c. *kuma=a mmja [ssjanaa-ssjana=gami=du ar]=ri.*
 this.place=TOP well RED-dirty=EMP=FOC COP=CNF
 ‘This place is rather dirty, eh?’ [phonological host: adjective]

Second, if quasi-quantifiers attach to argument phrases, they usually simply attach to case-marked noun phrases, but some (*=tjaaki* ‘only’, *=bakaar* ‘only’ and *=mai* ‘even; too’) must replace the nominative case marker. Third, *=tjaaki* ‘only’ and *=bakaar* ‘only’ can be placed either before or after the accusative *=ju* or dative *=n* (e.g. *uri=u=tjaakil ui=tjaaki=u* ‘it (ACC) only’; *ui=n=tjaaki/ui=tjaaki=n* ‘it (DAT) only’). Unlike the other quasi-quantifiers, these two usually come before a case marker, behaving like numerals in apposition with the head noun of a noun phrase (cf. (5)).

If we take distributional complexity as a scalar notion and if we consider that words involve a more complex distributional principle than a clitic (as argued by Zwicky 1985), we can point out that quasi-quantifiers are closer to words than case markers. In particular, the second and third characteristics noted above induce us to argue that *=tjaaki* and *=bakaar* might be better analysed as words (phonologically dependent words, as in (1)), just as numerals are treated as words.

Given that quasi-quantifiers attach syntactically to arguments, adjuncts and predi-

cates, the phonological host of a quasi-quantifier may vary widely. If the syntactic host is an argument phrase, the phonological host may be a case marker, (case-ellipted) nominal or numeral. The following examples illustrate the phonological hosts of =*mai* ‘too; even’ when attached syntactically to argument phrases (which are in brackets).

- (14) [*buuc=cu*]=*mai* *ibi-tigaa* *ibi-ru*.
 sugarcane=ACC=too plant-CND plant-IMP
 ‘Plant sugarcanes, too, if you (can) do so.’ [phonological host: case marker]
- (15) [*sinsii*]=*mai* [*siitu-mmi*]=*mai* *sjooka=u=ba* *as-ta-m*.
 teacher=too pupil-PL=too school.song=ACC=TOP do-PST-RLS
 ‘Both teachers and pupils sang school songs.’ [phonological host: noun]
- (16) [*uma=nu* *f-taar*]=*mai* *if-tar=ri*.
 that.place=GEN two-CLF.PRS=too go-PST=CNF
 ‘The two people there went, too, right?’ [phonological host: numeral]

If the syntactic host is an adjunct phrase, the phonological host may be an adverbial word or a converb that heads an adverbial clause that is syntactically embedded as an adjunct clause.

- (17) *uri=a* *umissi-f=mai* *njaan*.
 3SG=TOP interesting-ADV=too NEG
 ‘That’s not interesting, either.’ [phonological host: adverb]
- (18) [*vva=ga* *kuu-ba*]=*mai* *junuguu*.
 2SG=NOM come-CND=even same
 ‘Even if you come, (the result) would be the same.’ [phonological host: verb]

If the syntactic host is a predicate phrase, the phonological host may be the lexical verb element of a complex verbal predicate phrase (as was noted in (13a)) or the predicate nominal or adjective of a complex nominal/adjectival predicate phrase (as in (13b)).

3.4 Modal markers

Modal markers comprise the dubitative =*bjaam* ‘I wonder (if)’, another dubitative marker =*gagara* ‘I wonder (what, how, etc.)’, uncertainty =*paz* ‘maybe’, assertive =*dara*, emphatic =*doo* and reserved emphasis =*saa*. They all simply attach to a predicate phrase, either verbal or nominal (or adjectival, since a nominal predicate phrase may be headed by an adjective; see Section 3.3). For example, if the dubitative =*bjaam* is syntactically attached to a nominal predicate, the phonological host may be a noun, as in (19), or an adjective, as in (20).

- (19) *uri=a* *jamatu-pz̄tu=bjaam*.
 3SG=TOP mainland-person=I.wonder

- ‘I wonder if he is a Japanese mainlander.’ [phonological host: noun]
 (20) *nnama=gami=a sdass-sdas=bjaam=mi=ti=du.*
 now=EMP=TOP RED-cool=I.wonder=CNF=QT=FOC
 ‘(It’s) now cool, I wonder.’ [phonological host: adjective]

A verbal predicate may be headed by a verb or adjectival verb, so either of them can be the phonological host of a modal clitic.

- (21) *kan=nu=ru ur=bjaam=ti, muu=ju*
 crab=NOM=FOC exist=I.wonder=QT seaweed=ACC
ujukas-tigaa, bazakar-i-i=du u-tar.
 shake-CND raise.claw-INF-SEQ=FOC PROG-PST
 ‘(I thought) “I wonder if there is a crab,” and when I shake seaweed, (the crab) was raising its claw (to fight against my attack).’ [phonological host: verb]
 (22) *ban=jarru=mai aparagi-kar=bjaam=mi, uri=a=ju.*
 1SG=CPR=even beautiful-V.ADJ=I.wonder=CNF 3SG=TOP=COR
 ‘She’s perhaps more beautiful than I, I tell you.’ [phonological host: adjectival verb]

Modal markers may be phonologically attached to adverbs, as in (23B). This does not necessarily mean that the syntactic host includes an adjunct phrase as in the case of quasi-quantifiers; rather, it is more likely that an example like (23B) undergoes ellipsis of the predicate with which the adverb is associated. In (23), it is clear that the predicate *askutu* is ellipsed under identity.

- (23) A: *ssu-f as-kutu=ru?*
 red-ADV do-OBL=Q
 ‘Should we make (it) whitish?’
 B: *aai, aka-f=paz.*
 No red-ADV=maybe
 ‘No, (we should make it) red.’

3.5 Information-structure markers

Information-structure markers comprise focus markers (declarative =*du*, Yes-No interrogative =*ru* and WH interrogative =*ga*) and topic markers (object topic =*ba*, otherwise =*ja*). Except for the object topic =*ba*, which is necessarily attached to an object noun phrase, these markers occur on an argument phrase, adjunct phrase, the lexical part of a predicate phrase (complement, lexical verb or predicate nominal) and a restricted set of conjunctions (for focus markers only; e.g. *asii* ‘then’, *ttarjaa* ‘then’, *assuga* ‘but’ and *assiba* ‘so’). That is, they attach syntactically to almost any constituent of a sentence that can be focal or topical. In many cases an information-structure marker and a quasi-quantifier

tifier are paradigmatically selected, but when they do co-occur a quasi-quantifier must precede an information-structure marker.

The phonological host of an information-structure marker varies more widely than the other clitics reviewed so far, ranging from a noun to a case marker, verb, adjective (of any type), adverb and conjunction. The following examples illustrate the use of the declarative focus =*du*.

- (24) *uri=a pav=du a-tar.*
 3SG=TOP snake=FOC COP-PST
 ‘It was a snake.’ [phonological host: noun]
- (25) *pav=nu=du juu idi-i t-tar.*
 snake=NOM=FOC very exit-SEQ come-PST
 ‘SNAKES came out very (frequently).’ [phonological host: case marker]
- (26) *pav=nu juu=du idi-i t-tar.*
 snake=NOM very=FOC exit-SEQ come-PST
 ‘Snakes came out VERY (FREQUENTLY)’ [phonological host: adverb]
- (27) *pav=nu juu idi-i=du t-tar.*
 snake=NOM very exit-SEQ=FOC come-PST
 ‘Snakes CAME OUT very (frequently).’ [phonological host: verb]
- (28) *pav=nu idi-i uturus-ka-i-ba=du.*
 snake=NOM exit-SEQ come-V.ADJ-INF-CSL=FOC
 ‘Snakes came out and (were) FEARFUL, SO...’ [phonological host: inflected adjective]
- (29) *pav=nu uturus-uturus=du a-tar.*
 snake=NOM RED-fearful=FOC COP-PST
 ‘Snakes were FEARFUL.’ [phonological host: reduplicated adjective]
- (30) *pav=va uturus-munu=du a-tar.*
 snake=TOP fearful-NADJ=FOC COP-PST
 ‘The snakes were FEARFUL.’ [phonological host: nominal adjective]
- (31) *asii=du hira, pav=nu juu idi-i t-tar.*
 then=FOC you.know snake=NOM very exit-SEQ come-PST
 ‘THEN, you know, snakes came out very (frequently).’ [phonological host: conjunction]

3.6 Discourse markers

Discourse markers comprise =*ju* (corrective; =*Y* ‘(not X) but Y’), =*ca* (hearsay), =*da* (presentational; ‘how about...?’), =*ti(i)* (quotative), =*e(e)* (question) and =*i(i)* (confirmative, which functions like English tags). As indicated by bracketed segments, many have monomoraic and bimoraic variants.

The presentational =*da* is exceptional in that it is attached syntactically to a specific type of phrase: a topic-marked noun phrase (of any grammatical role), as in (32) to (34).

- (32) *kuma=a=da?*
 this.place=TOP=PRST
 ‘How about this place?’
- (33) *banti=ga buza=u=baa=da?*
 1PL=GEN uncle=ACC=TOP=PRST
 ‘How about (choosing) our uncle?’
- (34) *miz=zu fm=ma=da?*
 water=ACC get=TOP=PRST
 ‘How about getting water?’

The other discourse markers syntactically attach to any utterance unit, either sentence-medially or finally. As a result, there is no restriction with regard to the type of phonological host to which they attach. Whereas information-structure markers cannot attach phonologically (or syntactically) to adnominals or adnominal phrases, discourse markers can, as in (35).

- (35) A. *nza=nu pztu=kara=ga=gara=i.*
 what.place=GEN person=ABL=WH.FOC=I.wonder=CNF
 B. *kama=nu=ca.*
 that.place=GEN=HS
 ‘(That’s the person) of that place, I heard.’

(35B) must be an example of ellipsis, in which the word *psztu* ‘person’ in A’s utterance is ellipted under identity.

Discourse markers can also be attached syntactically and phonologically to a restricted set of conjunctions as well (e.g. *mmja* ‘well; you know; indeed’, *naugara* ‘well’, etc.). The following example illustrates the syntactic context in which the confirmative *=i(i)* occurs (note also that the hearsay marker *=ca* occurs in the sentence-initial adverbial and in the sentence-final predicate).

- (36) *nkjaan=du=ca, njkaan=du=i, kam=nu=i,*
 old.times=FOC=HS old.times=FOC=CNF god=NOM=CNF
doobuc=cu=i, icimus=su=i, mmna
 animal=ACC=CNF living.thing=ACC=CNF all
kam=nu mai=n zaa=nkai
 god=GEN front=DAT throne=ALL
acmar tukja=nu a-tar=ca.
 gather time=NOM exist-PST=HS
 ‘Once upon a time, (there) was a time when a god (ordered) animals, living things, all of them, (to) gather in front of his throne.’

4. Morphosyntactic aspects of Irabu clitics (2): Syntactic rules

The preceding section has revealed that the distributions of clitics are governed by simple principles which put clitics at the left edge of syntactic units (except for certain quasi-quantifiers; see Section 3.3), which make them syntactically special and distinguish them from words. On the other hand, it was also noted that clitics are distinct from affixes as well, since clitics are in construction with phrases or clauses, rather than with particular types of stems (see Table 2 in Section 3.1).

This section attempts to reveal another important characteristic of clitics that pertains to their special syntactic status: even though clitics occur at a syntactic level with their ability to be in construction with phrases, they are nevertheless partially immune to syntactic rules. I examine two syntactic rules: ellipsis (under identity) and movement. I illustrate how words are subject to these syntactic rules and how clitics are immune to one or both of them.

4.1 Ellipsis under identity

The ellipsis rule targets phrases. Words are subject to the rule as they are minimal phrases. Also, there is an indication that case markers are subject to the rule. I could not find any test that justify or falsify the claim that the rule targets other kinds of clitics.

If there is an overlapped part in two sentences (i.e. a part that has ‘anaphoric linkage’, in Zwicky’s 1985 terms), the ellipsis rule can operate to delete such a part. In the following dialogue, the utterance of B, which is a response to A, may be one of (a) to (c), and they can safely be considered to have undergone the ellipsis rule.⁴⁾

- (37) A. *nza=nu* *kookoo=nkai=ga* *if-tar?*
 what.place=GEN high.school=ALL=WH.FOC go-PST
 ‘Which high school did (you) go to?’
- (38) B. a. *ucnaa=nu* *kookoo=nkai.*
 Okinawa=GEN high.school
 ‘(I went) to a high school in Okinawa.’
 b. *ucnaa=nu* (*kookoo=nkai*).
 Okinawa=GEN high.school=ALL
 ‘(I went to a high school) in Okinawa.’
 c. **ucnaa=nu* (*kookoo*)=*nkai.*
 Okinawa=GEN high.school=ALL=FOC
 ‘(I went to (a high school) in Okinawa.’

As illustrated in (38c), case clitics cannot be stranded after ellipsis. This might appear to indicate that the ellipsis rule deletes a phrase, making it impossible for phrasal modifiers only to be deleted. However, this is misleading, as (38b) is possible. The ellipsis rule thus

targets the construct [noun + case clitic] as a whole. More precisely, since the noun *koo-koo* is a minimal noun phrase in itself, the deleted construct is [noun phrase + case clitic].

As was noted in Section 3.2, it is clear that a case clitic is attached to a phrase as a whole. The relative order of phrase formation, case clitic attachment and ellipsis is thus formulated as follows:

(39) Phrase formation → Case clitic attachment → Ellipsis rule

The ellipsis rule targets the [noun phrase + case clitic] as in the case of (38) above, but there is another ellipsis rule in which a noun phrase alone is deleted, giving rise to a stranded case clitic in the derived sentence. In Irabu, there is a headless relative clause structure, where a contextually recoverable (i.e. anaphorically linked) head noun may be deleted and the case marker is stranded, as in *fautar=ru mucikuu* (bring that which you have eaten), where the accusative =ru is stranded after the deletion of the head noun and the case marker is in effect directly attached to the adnominal clause. This phenomenon alone may allow two possible analyses with regard to the relative order of clitic attachment and ellipsis: (a) clitic attachment first and then ellipsis, and (b) the other way round. However, (a) alone can explain both cases of (38c) and the headless relative clause structure.

An example like (38c) is found in Japanese, called a “truncated possessive phrase” (Martin 1975), in which the head noun carrying a possessive modifier phrase undergoes the ellipsis rule and the head noun alone is deleted under identity.

- (40) A. *aitu=no kuruma=ga nusum-are-ta=tte.*
 3SG=GEN car=NOM steal-PASS-PST=HS
 ‘His car has been stolen, I heard.’
 B. *aitu=no=ga?*
 ‘His (car)?’

In Irabu, such a truncation is impossible, and the head noun must be left intact or replaced by a pro-form.

- (41) A. *kai=ga kuruma=nu=du nism-ai-tar=ca.*
 3SG=GEN car=NOM=FOC steal-PASS-PST=HS
 ‘His car has been stolen, I heard.’
 B. a. *kai=ga kuruma=nu=ru?*
 3SG=GEN car=NOM=YN.FOC
 ‘His car?’
 b. *kai=ga munu=nu=ru?*
 3SG=GEN thing=NOM=YN.FOC
 ‘His one?’

- c. **kai=ga=nu=ru?*
 3SG=GEN=NOM=YN.FOC
 ‘His (car)?’

As Vance (1993) points out, truncated possessive constructions in Japanese does not in itself tell anything with regard to the relative ordering of case marker attachment and ellipsis. A case marker may be attached to a noun phrase first then the ellipsis rule may delete the noun phrase alone, or a noun phrase may be deleted first and the case marker may then be attached (in effect directly to the possessor phrase). Either analysis predicts the same attested pattern like (40B). However, in the case of Irabu the latter analysis is impossible.

4.2 Movement

The movement rule in Irabu applies to phrases. In effect, they can apply to words since words are able to constitute minimal phrases on their own. Let us illustrate this with the following example, in which there are two noun phrases and one predicate phrase, each indicated by square brackets.

- (42) [*unu asb-i-ur* *jarabi-mmi*]=*nkai* [*mm=nu* *zz*]=*zu*
 that play-INF-PROG child-PL=ALL potato=GEN rice.ball=ACC
 [*fii-kutu*]=*dara*
 give-OBL=ASR
 ‘(We) have to give potato balls to those kids who are playing.’

The first noun phrase is composed of two modifiers: the adnominal word *unu* ‘that’ and the verb *asbiur* ‘be playing’, which constitute minimal adnominal phrases each, modifying the head noun *jarabimmi* ‘children’. This entire phrase is cliticised by the allative case marker =*nkai*. The second noun phrase is composed of the noun *mm* ‘potato’ (as a minimal noun phrase that serves as a modifier of a superordinate noun phrase in brackets) and the head noun *zz* ‘rice ball’, to which the accusative case marker is attached. The predicate phrase is composed of the head verb alone, to which the modal marker =*dara* (assertive) is attached phonologically.

In the first noun phrase, the adnominal word and verb can be interchangeable as they constitute minimal phrases within the noun phrase. By contrast, the head noun of the phrase cannot be moved without moving these modifiers, as the movement rule applies to the entire phrase. The same holds true for the second noun phrase, where it is impossible to move the head alone. Observe (43) below, where the first and second noun phrases in (42) are interchanged.

- (43) [*mm=nu* *zz*]=*zu* [*unu asb-i-ur* *jarabi-mmi*]=*nkai*
 potato=GEN rice.ball=ACC that play-INF-PROG child-PL=ALL

fii-kutu]=dara.
 give-OBL=ASR
 ‘(We) have to give potato balls to those kids who are playing.’

The predicate verb can also be moved (though it is a highly marked structure as Irabu strongly prefers to put predicates sentence-finally), as it constitutes a minimal predicate phrase.

(44) [*mm=nu zz]=zu [*fii-kutu]=dara [*unuasb-i-ur*
 potato=GEN rice.ball=ACC give-OBL=ASR that play-INF-PROG
 jarabi-mmi]=nkai
 child-PL=ALL
 ‘(We) have to give potato balls to those kids who are playing.’**

It is noted that it is impossible for a clitic (or affix) alone to be moved, indicating that clitics and affixes cannot constitute minimal phrases on their own. Also noted is the fact that where a phrase is moved, the clitic that attaches to the phrase must also occur together with the phrases that have been moved. For example, in (43) the second noun phrase of (42) is moved to sentence-initial position, still carrying the accusative marker together. It is impossible for the accusative case alone to be left at the original position. In short, not only are case clitics immune to the movement rule on their own right but they have to cling to the noun phrase when it is subject to a movement rule. This holds true for other clitics such as quasi-quantifiers, information structure markers and modal markers, all of which attach to phrases.

There are two possibilities regarding the relative order of clitic attachment and movement: clitics may be attached before the movement rule applies (i.e. the entire construct of a phrase plus a clitic is moved), or they may be attached after the movement rule applies. The surface fact is not contradictory to either possibility.

There is one argument in favour of the latter analysis, though it is restricted to the accusative case marker alone. In Irabu, there are two accusative case markers, default accusative =*ju* (ACC) and second accusative =*ja* (ACC2). The second accusative designates low transitivity, one syntactic indication of which is reflected in the fact that the object marked with the second accusative and the verb that follows are always adjacent (Shimoji 2008). If an element breaks up the bond of the O marked with the second accusative and the verb ($O_{[ACC2]}V \rightarrow O_{[ACC1]}[X]V$), the default accusative occurs. For example, if the adjunct *ic=mai* ‘always’ of (45) intervenes the object marked with the second accusative and the verb, we get something like (46) below, with ACC2 obligatorily replaced by the default accusative.

(45) *ic=mai asi=a kak-i-i=du niv-vi-ur.*
 always=too sweat=ACC2 have-INF-SEQ=FOC sleep-INF-PROG

- ‘(S/he) is sleeping, having sweat all the time.’
- (46) *asi=u* *ic=mai* *kak-i-i=du* *niv-vi-ur*.
 sweat=ACC always=too have-INF-SEQ=FOC sleep-INF-PROG
 ‘(S/he) is sleeping having sweat all the time.’

The fact that the accusative case marking is sensitive to the re-ordering (i.e. movement) of the object noun phrase and the predicate tells us that the case marking may occur after the movement rule applies. If we assume that what applies to the accusative case is systematically applicable to other case markers, then it is possible to speculate that all case clitics attach after the movement rule applies. Furthermore, since all other clitics occur after case markers if they co-occur, we might further speculate that these clitics also occur after the movement rule applies.

4.3 Summary

The examination of the ellipsis rule indicates that a case clitic can be a target of ellipsis, as the rule deletes [noun phrase + case clitic] (as in cases like (38)). On the other hand, the headless relative clause formation deletes a noun phrase alone, with the case clitic being stranded after the ellipsis. To explain both cases satisfactorily, case clitic attachment must precede the ellipsis rule. On the other hand, there is no clear indication that the other clitics also occur before the ellipsis rule applies (but see below for discourse markers).

With regard to the movement rule, the surface facts tell us two different analyses with regard to the relative order of clitic attachment and movement. On the one hand, it is possible for a clitic to be attached and then the movement rule applies; on the other, it is not contradictory either if a clitic is attached after the movement rule applies, since what we observe at a surface level is a (moved) phrase with a case clitic. There is one strong indication that the accusative case marking occurs after the movement rule applies, from which we can draw a reasonable guess that all case clitics behave like the accusative case clitic, occurring after the movement rule applies. Based on the fact that all other kinds of clitics follow case clitics in clitic chains, the further speculation is made that they also occur after the movement rule applies. Thus, I weakly argue that clitic attachment follow the movement rule applies, with the speculation that clitics are immune to the movement rule at all.

In sum, it is impossible to state clearly that all clitics in Irapu are immune to syntactic rules, as theoretical models such as Zwicky’s (1985) predict for the clitic as a theoretical construct. Zwicky’s model predicts that clitics should occur after the application of all syntactic rules like ellipsis under identity and movement target syntactic constituents (including words). Based on this model, Irapu case markers are like words with respect to the ellipsis rule, but can be considered to be a good example of clitic as far as the ellipsis rule is concerned.

Discourse markers are distinct from the other clitics, since the former are attached to any utterances rather than specific types of phrases (Section 3.6). Here, the surface fact

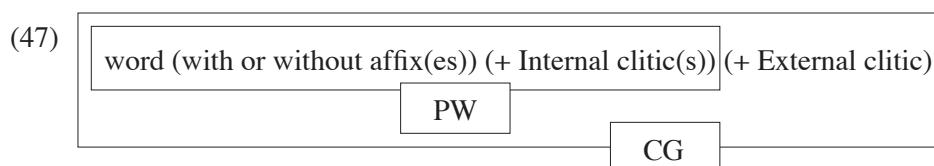
tells us that discourse markers can occur after syntactic rules such as movement and ellipsis apply. Whereas a phrase is a structural unit and there are reasons to believe that they are subject to syntactic rules such as the ellipsis rule and movement rule, an utterance is a surface unit and must have undergone all syntactic processes including movement and ellipsis under identity. This theoretical assumption leads to the analysis that discourse markers must thus be attached after all these processes apply. As will be demonstrated in Section 6, there are phonological indications that half of the discourse markers (three out of six) attach after all major phonological processes apply, which also supports the view that discourse markers occur at a surface or near-surface level.

5. Phonological aspects of clitics

This and following sections examine the phonological characteristics of clitics. Clitics and affixes are both phonologically dependent, and there is no need to distinguish between clitics and affixes in this regard. However, in Section 6 it will be demonstrated that three out of seven discourse markers are phonologically less dependent, and we need to distinguish between such clitics on the one hand (External clitics) and the other clitics (Internal clitics) and affixes on the other.

5.1 Phonological word and clitic group

Before going further, let us introduce the units Phonological Word (PW) and Clitic Group (CG) for the ease of the subsequent discussion (see the schematic structure in (47) below).



A PW is the phonologically coherent domain, and the phonological independence in this paper is judged by whether an element can stand as a PW. A PW usually consists of a word (with or without affixes) and a set of clitics that are fully integrated into the host PW. I call such clitics Internal clitics.

There is another set of clitics, External clitics, which are literally external to PWs, and are treated as a sort of “extra” elements, even though they cannot stand as PWs on their own either. An PW plus an External clitic is called a CG, as suggested by Selkirk (1984) and Nespor and Vogel (1986). The PW is defined as an independent domain of foot-based alternating rhythm of tone features (or “HL alternation,” Shimoji 2009a). HL alternation comprises two ordered processes: foot building and tone assignment.

5.2 Foot building

A foot is either bimoraic (default) or trimoraic. Footing goes from left to right iteratively. A stray mora resulting from either of the following is included in the preceding foot.

(48) Bimoraic footing on a PW with an odd number of moras ($PW_n = PW$ of n morae) results in one stray finally, which is included in the final foot.

- | | | |
|-------------|-------------------------------|------------------|
| a. PW_2 : | <i>pana</i> ‘nose’ | (pana) |
| b. PW_3 : | <i>katana</i> ‘knife’ | (katana) |
| c. PW_4 : | <i>utugaja</i> ‘jaw’ | (utu)(gaja) |
| d. PW_5 : | <i>banckira</i> ‘guava’ | (ban)(ckira) |
| e. PW_6 : | <i>koozaburoo</i> ‘Kozaburo’ | (koo)(zabu)(roo) |
| f. PW_7 : | <i>oostoraria</i> ‘Australia’ | (oo)(sto)(raria) |

(49) Polymoraic morphs always commence a foot, i.e. the left boundary of a polymoraic morph always coincides with the left boundary of a foot.

- | | |
|------------------------------|------------------------------|
| a. <i>katana-nagi</i> | b. <i>katana=mai</i> |
| knife-and.so.on | knife=too |
| ‘knife, and so on’ | ‘knife, too’ |
| i. (katana)(nagi) | i. (katana)(mai) |
| *ii. (kata)na(nagi) | *ii. (kata)na(nagi) |
| c. <i>oostoraria-nagi</i> | d. <i>oostoraria=mai</i> |
| Australia-and.so.on | Australia=too |
| ‘Australia, and so on’ | ‘Australia, too’ |
| i. (oo)(sto)(raria)(nagi) | i. (oo)(sto)(raria)(mai) |
| *ii. (oo)(sto)(rari)a (nagi) | *ii. (oo) (sto)(rari)a (mai) |

On the other hand, monomoraic morphs are simply part of the preceding host, to which the default footing applies.

- | | |
|--------------------------|-------------------------|
| (50) a. <i>katana=nu</i> | b. <i>oostoraria=nu</i> |
| knife=NOM | Australia=NOM |
| ‘knife:nom’ | ‘Australia:nom’ |
| (kata)(nanu) | (oo)(sto)(rari)(anu) |

A sequence of monomoraic morphs are also part of the preceding host to which the default footing applies. Compare (51a) with (49a, b), (51b) with (49c, d).

- | | |
|-----------------------------|----------------------------|
| (51) a. <i>katana=nu=du</i> | b. <i>oostoraria=nu=du</i> |
| knife=NOM=FOC | Australia=NOM=FOC |
| ‘knife:nom:foc’ | ‘Australia:nom:foc’ |
| (kata)(nanudu) | (oo)(sto)(rari)(anudu) |

5.3 Tone assignment

Based on the pre-existing foot structure, tone is assigned according to the following principles. First, as shown in the left side of (52) below, adjacent feet (F) are grouped into “foot groups” (indicated by \square below), which can consist of up to three feet. Thus, if a foot group consisting of three feet is going to contain one more foot, it must be regrouped into two foot groups (b). Second, /H/ is assigned to the left-most foot of a foot group. The other feet are left unmarked (/Ø/), which is realised as a lower pitch.

(52) Foot group formation		Tone assignment
a. \square (F)	>	\square (H)
b. \square (F) + (F) = \square (F)(F)	>	\square (H)(Ø)
c. \square (F)(F) + (F) = \square (F)(F)(F)	>	\square (H)(Ø)(Ø)
d. \square (F)(F)(F) + (F) = \square (F)(F) \square (F)(F)	>	\square (H)(Ø) \square (H)(Ø)
e. * \square (F)(F)(F)(F)	>	* \square (H)(Ø)(Ø)(Ø)
f. \square (F) + (F) = * \square (F) \square (F)	>	* \square (H) \square (H) (cf. (b))

(53)	Foot group structure	Rhythmic structure
PW ₂ : <i>tur</i> ‘bird’	\square (tur)	\square (H)
PW ₃ : <i>tur=du</i>	\square (turdu)	\square (H)
PW ₄ : <i>tur-gama</i>	\square (tur)(gama)	\square (H)(Ø)
PW ₅ : <i>tur-gama=du</i>	\square (tur)(gamadu)	\square (H)(Ø)
PW ₆ : <i>tur-gama-mmi</i>	\square (tur)(gama)(mmi)	\square (H)(Ø)(Ø)
PW ₇ : <i>tur-gama-mmi=du</i>	\square (tur)(gama)(mmidu)	\square (H)(Ø)(Ø)
PW ₈ : <i>tur-gama-mmi=kara</i>	\square (tur)(gama) \square (mmi)(kara)	\square (H)(Ø) \square (H)(Ø)
PW ₉ : <i>tur-gama-mmi=kara=du</i>	\square (tur)(gama) \square (mmi)(karadu)	\square (H)(Ø) \square (H)(Ø)
PW ₁₀ : <i>tur-gama-mmi=kara=gami</i>	\square (tur)(gama) \square (mmi)(kara)(gami)	\square (H)(Ø) \square (H)(Ø)(Ø)

Note: =*du* (focus), -*gama* (diminutive), -*mmi* (plural), =*kara* (ablative), =*gami* (emphasis)

The following pair of examples only differ in the stem to which they attach. In (54a), the stem has one foot, whereas in (54b) the stem has two feet. Since the total number of feet differs in these two word-pluses, rhythmic structure also differs.

(54) a. <i>midum-gama-mmi=kara</i> woman-DIM-PL=ABL ‘from girls’ (midum)(gama)(mmi)(kara) (H)(Ø)(H)(Ø)	b. <i>bikidum-gama-mmi=kara</i> man-DIM-PL=ABL ‘from boys’ (biki)(dum)(gama)(mmi)(kara) (H)(Ø)(H)(Ø)(Ø)
--	---

The first /H/ in a series within a PW is phonetically realised as the highest pitch, allowing us to distinguish between PW₈ and two PW₄ in a row, for example. It follows from the above principle of HL alternation that a sequence of two /H/ feet indicates the termination of a prosodic domain between the two /H/ feet, as in (55), and that a sequence of two toneless feet indicates the termination of a prosodic domain after them, as in (56). Thus, the rhythmic alternation serves as a kind of boundary marker, which marks off each PW.

- (55) *ami=nu* *f-fi-u-m.*
rain=NOM fall-INF-PROG-RLS
‘It’s raining.’
(aminu) (ffi)(um)
(H) (H)(Ø)
- (56) *uttu-ssu-mmi=nu* *cn-gama*
sibling-younger-PL=GEN clothes-DIM
‘younger siblings’ clothes’
(uttu)(ssu)(mminu) (cn)(gama).
(H)(Ø)(Ø) (H)(Ø)

5.4 Summary

Since HL alternation requires its domain to be footed before tone is assigned, the domain must consist of one or more feet, to which the tone assignment rule applies. Thus, for an element to be phonologically independent, it must satisfy the minimality constraint (which says that it must have at least one foot, or two morae) *and* it must also serve as an independent domain in which the tone assignment occurs. Therefore, if an element is monomoraic, it is necessarily phonologically dependent. If an element is bimoraic or longer, then the next question is whether it serves as an independent domain of tone assignment. Only words are phonologically independent. Clitics and affixes are phonologically dependent either because they are monomoraic (e.g. the nominative case marker =*ga*) or because they cannot be an independent domain of tone assignment even though they are polymoraic (e.g. the ablative case marker =*kara*).

A closer examination of the foot building and tone assignment points to an interesting discrepancy between many discourse markers (which are analysed as External clitics) and the other dependent elements (Internal clitics and affixes). Internal clitics and affixes are dependent in terms of foot building and/or tone assignment. That is, they do not necessarily come at the left edge of the footing domain, nor do they serve as an independent domain for tone assignment. In Section 5.2, it was noted that foot building does not refer to the distinction between (Internal) clitics and affixes, to the extent that any monomoraic morph is treated as a part of the preceding host for the purpose of footing, and any polymoraic morph is treated as a distinct domain of footing. Internal clitics and affixes show an identical behaviour in tone assignment as well, in that they are always treated as internal members of a PW.

By contrast, as will be discussed in the following section, External clitics behave differently from words on the one hand, and from Internal clitics and affixes on the other. They differ from words in that they do not constitute PWs on their own. They also differ from Internal clitics and affixes in that the former are never integrated into the host PW.

6. External clitics

Out of the six discourse markers I examined in Section 3.6, the following three are analysed as External clitics: the corrective =*ju* (corrective; A=*ju* ‘(not B) but A’), hearsay =*ca* and confirmative =*i(i)* (which functions to confirm hearer’s sustained attention, like English tags). The confirmative may be lengthened (e.g. =*i* → =*ii*), which indicates that its phonological dependency fluctuates between the more phonologically dependent monomoraic form (which violates the minimality constraint; see Section 5.4) and more independent bimoraic form (which satisfies the minimality constraint).

6.1 Corrective =*ju* and hearsay =*ca*

The discourse particles =*ju* and =*ca* (hearsay) are identified as External clitics because of their peculiar behaviours of foot building, where they do not become a part of the preceding host for footing. This becomes clear when they are compared with Internal clitics and affixes, as in (57) and (58).

(57)	Underlying structure	Foot structure	Tone assignment	Pitch shape of PW
	a. <i>vva</i> ‘2SG’ + = <i>ga</i> (NOM)	(<i>vvaga</i>)	(H)	[HHH]
	b. <i>vva</i> ‘2SG’ + = <i>du</i> (PL)	(<i>vvadu</i>)	(H)	[HHH]
(58)	Underlying structure	Foot structure	Tone assignment	Pitch shape of PW
	a. <i>imsja</i> ‘fisherman’ + = <i>nu</i> (NOM)	(<i>im</i>)(<i>sjanu</i>)	(H)(Ø)	[HHLL]
	b. <i>imsja</i> ‘fisherman’ + = <i>ta</i> (PL)	(<i>im</i>)(<i>sjata</i>)	(H)(Ø)	[HHLL]

In (57), the second person pronoun *vva* takes the nominative case particle =*ga* (which is an Internal clitic) and the plural suffix =*du*. In (58), the noun *imsja* ‘fisherman’ takes the nominative case particle =*nu* (which is again an Internal clitic) and the plural suffix =*ta*. As is usual in bound monomoraic morphs, these bound morphemes are integrated in the host for the purpose of footing, yielding a trimoraic foot in (57) and two bimoraic feet in (58). Then tone is assigned according to the existing foot structure.

The situation is different when the same hosts are followed by External clitics. In (59), the second person pronoun *vva* is cliticised by the External clitics =*ju* and =*ca* both being outside of the footing domain, thus causing a bimoraic foot in each case.

(59)	Underlying structure	Foot structure
	a. <i>vva</i> ‘2SG’ + = <i>ju</i> (COR)	(<i>vva</i>) <i>ju</i>
	b. <i>vva</i> ‘2SG’ + = <i>ca</i> (HS)	(<i>vva</i>) <i>ca</i>

Tone is assigned by rule, as shown in (60) below. The H-toned foot in each example is realised as [HL] rather than the expected [HH], and the External clitic that attaches to the host is pronounced with a slightly higher pitch than the [L] of the host, or with a rise-fall contour (and phonetically with a slightly longer duration). The tonal characteristic of the External clitic is independent of the HL alternation, and is determined by the sentential intonation.

(60)	Underlying structure	Foot structure	Tone assignment	Pitch shape of PW
	a. <i>vva</i> ‘2SG’ + = <i>ju</i> (COR)	(vva) ju	(H) ju	[HL]
	b. <i>vva</i> ‘2SG’ + = <i>ca</i> (HS)	(vva) ca	(H) ca	[HL]

In (61), the noun *imsja* ‘fisherman’ is cliticised by the External clitics =*ju* (corrective) and =*ca* (hearsay).

(61)	Underlying structure	Foot structure	Tone assignment	Pitch shape of PW
	a. <i>imsja</i> ‘fisherman’ + = <i>ju</i> (COR)	(imsja) ju	(H)	[HHL]
	b. <i>imsja</i> ‘fisherman’ + = <i>ca</i> (HS)	(imsja) ca	(H)	[HHL]

Again, both clitics are outside of the footing domain, yielding a trimoraic foot in each case rather than two bimoraic feet as in the case of (58). The H-toned foot in each case of (61) is phonetically realised as [HHL] with the initial and final moras being lowered, and the External clitics are pronounced with a slightly higher pitch than the [L] of the host.

When the last foot of the host is /Ø/ as in (62) below, External clitics do not affect the pitch of the final foot of the host. Still, they are pronounced with a higher pitch than the final foot of the host, just as in the case of the above examples, demonstrating that their pitch realisation is independent of the tone assignment of the host.

(62)	Underlying structure	Foot structure	Tone assignment	Pitch shape of PW
	a. <i>aparagi</i> ‘beautiful’ + = <i>ju</i> (COR)	(apa)(ragi)ju	(H)(Ø)	[HHLL]
	b. <i>aparagi</i> ‘beautiful’ + = <i>ca</i> (HS)	(apa)(ragi)ca	(H)(Ø)	[HHLL]

The phonological behaviours of the discourse markers allow us to assume that the corrective =*ju* and hearsay =*ca* are attached after the footing and tone assignment apply to the host, i.e. after a PW is formed. That is, these clitics occur at the level of what we call CG (Clitic Group). The pitch lowering of the final mora of the host is a rule that applies to the CG. In summary, (63) is the ordering of rules that apply to a PW and a CG with regard to the attachment of the External clitics =*ju* and =*ca* (EC below).

(63) Reformulation of external clitic attachment

Underlying structure	Footing	Tone assignment	EC attachment	Surface output
a. <i>vva</i> '2SG'	→ (vva)	→ (H)	→ (H)=ju	[HL]=ju
b. <i>vva</i> '2SG'	→ (vva)	→ (H)	→ (H)=ca	[HL]=ca

PW

CG

6.2 Confirmative =i

Like the corrective =*ju* and hearsay =*ca* above, the confirmative =*i* is not integrated into the host in terms of footing. This is clearly seen in (64b), where the surface form *imsja=i* (the host *midum* + the clitic =*i*) does not constitute two-foot structure *(*im*)(*sjai*).

(64) Attachment of =*i*

Underlying structure	Footing	Tone assignment	EC attachment	Surface output
a. <i>vva</i> '2SG'	→ (vva)	→ (H)	→ (H)=i	[HH]=i
b. <i>imsja</i> 'fisherman'	→ (imsja)	→ (H)	→ (H)=i	[HHH]=i
c. <i>aparagi</i> 'beautiful'	→ (apa)(ragi)	→ (H)(Ø)	→ (H)(Ø)=i	[HHLL]=i

Unlike the other two discourse particles, the confirmative =*i* never causes the final mora of the host to be lowered, which indicates that =*i* is more independent in terms of prosody than the other two particles, since there is no prosodic interaction between the host and the clitic. This observation is further supported by another fact that the confirmative =*i* is often realised as the bimoraic =*ii*, and it can be used as a complete utterance without any preceding element (e.g. as a short response to the addressee).

6.3 Clitic group and morpho-phonological rule

The existence of the CG domain as against the PW domain is further reinforced by the availability of a morphophonological rule that only applies within the PW domain: the Geminate Copy Insertion Rule. The Geminate Copy Insertion rule checks the phonotactic well-formedness of a PW, fixing the ill-formed phonotactic pattern 'C.(G)V' (a coda directly followed by an onset-less nucleus) within the PW domain. Since Internal clitics occur within a PW, they are subject to the Geminate Copy Insertion when they begin in an onset-less syllable (e.g. topic =*ja*, accusative =*ju*, etc.) are attached to hosts that end in a coda consonant. Let us examine the attachment of the accusative =*ju* to different hosts that end in different consonants.⁵⁾

- (65) a. *kan* 'crab' + =*ju* (ACC) → *kan=nu*
 b. *kam* 'god' + =*ju* (ACC) → *kam=mu*
 c. *pav* 'snake' + =*ju* (ACC) → *pav=vu*

- d. *paz* ‘fly’ + =*ju* (ACC) → *paz=zu*
 e. *par* ‘needle’ + =*ju* (ACC) → *par=ru*
 f. *pus* ‘star’ + =*ju* (ACC) → *pus=su*
 g. *pac* ‘bee’ + =*ju* (ACC) → *pac=cu*

By contrast, the Geminate Copy Insertion does not operate on a CG. Let us examine the attachment of the corrective =*ju* to the same set of hosts.

- (66) a. *kan* ‘crab’ + =*ju* (COR) → *kan=ju* [kanⁿju] ~ [kaŋ.ju]
 b. *kam* ‘god’ + =*ju* (COR) → *kam=ju* [kam^mju]
 c. *par* ‘needle’ + =*ju* (COR) → *par=ju* [pa||ju]
 d. *pav* ‘snake’ + =*ju* (COR) → *pav=ju* [pavju]
 e. *paz* ‘fly’ + =*ju* (COR) → *paz=ju* [paʒju]
 f. *pus* ‘star’ + =*ju* (COR) → *pus=ju* [pus(i)ju]
 g. *pac* ‘bee’ + =*ju* (COR) → *pac=ju* [pats(i)ju]

Phonetically speaking, in (66a-c) where the host-final consonants are resonants, they are copied onto the glide /j/ of =*ju* as off-glides. These phonetic realisations might first appear to be a result of the Geminate Copy Insertion, inducing us to give phonemic representations *kan=nju*, *kam=mju* and *par=rju*. However, there is no special reason to consider these off-glides as phonemic. In fact, even if a special rule is absent, it is almost impossible to pronounce the sequence of the host-final consonant and the clitic-initial /j/ without any off-glide. For example, (66b) could be pronounced as [kam.ju] but this is possible only if the host and clitic are pronounced as two distinct utterance units. Thus, the off-glide is completely predictable from the combination of a resonant coda and the clitic-initial /j/.

The non-phonemic status of the off-glides is most clear in (66a). It may be realised as [kannⁿu], but may also be realised as [kaŋ.ju] (or [kaŋ.ju]), with no off-glide nasal. This phonetic difference directly comes from the difference in the allophonic realisation of /n/. The utterance-final allophony of the phoneme /n/ (e.g. *kan* ‘crab’) varies, ranging from the underlying [n] to [ŋ] (or [N]): *kan* ‘crab’ [kan] ~ [kaŋ]. The fact that [kannju] carries the geminated onset on [u] is explained by the fact that the articulatory gesture of the coronal allophone of /n/ toward the following /j/ results in a transitional off-glide. This transitional off-glide is absent in the dorsal allophone of /n/ followed by /j/. Thus, the off-glide is completely determined by the allophonic realisation of the host, and there is no phonemic ground on which to argue that the Geminate Copy Insertion occurs on the combination between the corrective =*ju* and its host.

Exactly the same argument holds for the confirmative =*i*. In (67a), the phonetic realisation may be either [kanni] if the allophone of the host-final /n/ is [n] or [kaŋ.ju] (or [kaŋ.ju]) if the allophone of the host-final /n/ is [ŋ] or [N].

(67) a.	<i>kan</i> ‘crab’	+ = <i>i</i> (CNF)	→ <i>kan=i</i>	[kan ⁿ i] ~ [kaŋ.i]
b.	<i>kam</i> ‘god’	+ = <i>i</i> (CNF)	→ <i>kam=i</i>	[kam ^m i]
c.	<i>par</i> ‘needle’	+ = <i>i</i> (CNF)	→ <i>par=i</i>	[pa ^l i]
d.	<i>pav</i> ‘snake’	+ = <i>i</i> (CNF)	→ <i>pav=i</i>	[pavi]
e.	<i>paz</i> ‘fly’	+ = <i>i</i> (CNF)	→ <i>paz=i</i>	[pazi]
f.	<i>pus</i> ‘star’	+ = <i>i</i> (CNF)	→ <i>pus=i</i>	[pusii]
g.	<i>pac</i> ‘bee’	+ = <i>i</i> (CNF)	→ <i>pac=i</i>	[patsii]

In summary, based on the fact that the Geminate Copy Insertion rule does not operate on the combinations between hosts and External clitics, I conclude that this morpho-phonological rule is not applicable in the CG domain.

7. Concluding remarks

This paper examined the morphosyntactic and phonological properties of clitics in Irabu. Morphosyntactically, they are word-like in the sense that they occur at phrasal or clausal level, combining with phrases or clauses. The construction with phrases and clauses makes clitics distinguished from affixes, which are affixed to stems. On the other hand, they are syntactically not word-like in that they are governed by fairly simple distributional principles and are partially immune to syntactic rules, especially the movement rule. Phonologically, it was argued that a distinction should be made between Internal and External clitics.

By observing both morphosyntactic and phonological properties of clitics independently, I pointed to an inter-relationship between the morphosyntactic property of discourse markers and their phonological behaviours. The distributional characteristics of discourse markers indicate that they occur utterance-finally, that is, they attach at a near-surface syntactic level (after all syntactic rules apply). This is independently confirmed by the phonological fact that half of the discourse markers attach after PWs are formed.

The definition and identification of clitics in this paper is pre-theoretical in nature, treating many dependent elements as clitics even though they might otherwise be analysed as words in other theoretical persuasions. Based on Zwicky and Pullum (1983) and Zwicky (1985), Vance concludes that there is no strong indication that Japanese particles (elements that are similar in function with ones we discussed in this paper) are clitics, arguing in favour of an analysis that most of them are indeed words. He clearly argues against the idea that Japanese particles are affixes, but he is more careful in examining whether they are words or clitics. Irabu particles (case markers, quasi-quantifiers, information-structure markers, modal markers, and discourse markers) are clearly not affixes since they are in construction with phrases. On the other hand, just as in the case of Japanese, it was a tricky task to identify an element as a clitic or word in terms of morphosyntactic criteria, especially of the applicability of syntactic rules. It has been demonstrated that half the discourse markers, which are claimed to occur after all syntactic rules apply

and occur at a near-surface level in the phonological sense, are clear examples of the clitic in the stricter sense with which Zwicky's model defines it as a theoretical construct. The other clitics are not readily justified as clitics in Zwicky's terms. In particular, case markers are clearly subject to a syntactic rule (ellipsis rule), which would allow us to regard them as words; however, the applicability of another syntactic rule (movement rule) tells that they are immune to this syntactic rule and they can occur after the rule applies, thus making us believe that they are clitics in Zwicky's theoretical prediction.

Notes

- 1) The etymologically related form *-dus* (as in *mii-dus-tar* 'looked' (look-V.FOC-PST)) is synchronically a single morpheme (verbal focus suffix). See Shimoji 2010 for detail.
- 2) Admittedly, treating these category-specific elements as 'post-inflectional affixes' is a problem if we take inflection as a morphological process that closes off the word formation. However, treating such elements instead as 'clitics' simply because of their occurrence after inflection is more problematic, since this treatment gives a fatal result to the characterisation of clitics, by admitting the category-specific property for clitics, a property that should be best handled in morphology rather than syntax.
- 3) A given adjectival root (*taka-* 'high', *sabic-* 'lonely', etc.) may be transformed into a nominal, verbal or adjective (see Shimoji 2009b for detail): Adjectival nouns are a subclass of the nominal (e.g. *taka-munu* 'high', *sabic-munu* 'lonely', etc.) and they behave exactly like ordinary nouns, even though their meaning restricts their function (i.e. they are not normally used as arguments, though they can). Adjectival verbs constitute a subclass of the verbal (e.g. *taka-ka-tar* 'was high', *sabic-ka-tar* 'was lonely', etc.), as their inflectional morphology and syntactic distribution are exactly like those of ordinary verbs. Adjectives are distinct both from nominals and verbals, which are defined as reduplicated forms of adjectival roots (*taka-* 'high' → *takaa-taka* 'high', etc.).
- 4) It is also possible to respond like *ucnaa=nu kookoo* 'A high school in Okinawa', which consists of a noun phrase alone. This might first appear to indicate that the underlying *ucnaa=nu kookoo=nkai* undergoes the deletion of *=nkai* under identity. However, it is also perfectly reasonable to assume that it is not really related to (37) but a newly introduced copular sentence without an overt subject noun phrase, as in (it was) a high school in Okinawa. That is, the noun phrase here may not be a result of ellipsis. In fact, it is possible to add a copula *atar* '(it) was'. Hence, it is impossible to determine whether this stranded noun phrase is really an example of ellipsis, and we should exclude it in the subsequent discussion.
- 5) The analysis that the accusative marker is *=ju* is controversial: one can alternatively analyse the underlying form of the marker as *=u*, as I actually did in my grammar of Irabu (Shimoji 2008). In this analysis, the comparison with the corrective *=ju* in (66) may not make much sense, as the two are different in form, and the morphological difference may result from this difference and not from the difference in the target domain of the Geminate Copy Insertion. However, in this case the comparison with the confirmative *=i* does make a perfect sense, as in (67), since both (i.e. the accusative *=u* and the confirmative *=i*) begin in a vowel. Either analysis, therefore, points to the difference in the availability of Geminate Copy Insertion rule between the accusative case clitic and a discourse marker clitic (corrective or confirmative).

References

- Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Anderson, Stephen R. 2005. *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.
- Bickel, Balthasar and Johanna Nichols. 2007. Inflectional Morphology. In Shopen, Timothy, ed., *Language typology and syntactic description (III)*, Cambridge: Cambridge University Press. Second edition.
- Carstairs, Andrew. 1981. *Notes on affixes, clitics, and paradigms*. Bloomington, IN: Indiana University Linguistics Club.

- Dixon, R.M.W., and Alexandra Y. Aikhenvald. 2002. *Word: a typological framework*. In Dixon, R.M.W., and Alexandra Y. Aikhenvald, eds, *Word: a cross-linguistic typology*, 1–41, Cambridge: Cambridge University Press.
- Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45 (1): 31–80.
- Johannessen, Janne Bondi. 1998. *Coordination*. New York: Oxford University Press.
- Karimata, Shigehisa. 2005. Okinawaken Miyakojima Hiraahooogen no foneemu nit suite [Phonemes in the Hirara dialect of Miyako Island, Okinawa Prefecture]. *Nihon Tooyoobunka Ronshuu* 11: 67–113.
- Klavans, Judith L. 1985. The Independence of Syntax and Phonology in Cliticization. *Language* 61: 95–120.
- Martin, Samuel E. 1975. *A Reference Grammar of Japanese*. New Haven, CT: Yale University Press.
- Matthews, Peter. 1974. *Morphology*. Cambridge: Cambridge University Press.
- Miller, Philip H. 1992. *Clitics and constituents in phrase structure grammar*. New York: Garland.
- Nespor, Marina, and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris Publications.
- Pellard, Thomas. 2006. Miyakoshohooogen no onin no mondaiten [Problems in the phonology of Miyako dialects]. Paper read at Study Group on Ryukyuan Languages, Kyoto University.
- Selkirk, Elisabeth. 1984. *Phonology and syntax: the relation between sound and structure*. Cambridge, MA: MIT Press.
- Shimoji, Michinori. 2008. *A grammar of Irabu, a Southern Ryukyuan language*. PhD dissertation. ANU.
- Shimoji, Michinori. 2009a. Foot and rhythmic structure in Irabu Ryukyuan. *Gengo Kenkyū* 135: 87–113.
- Shimoji, Michinori. 2009b. The adjective class in Irabu Ryukyuan. *Studies in the Japanese Language (Nihongo no Kenkyū)* 5 (3): 22–40.
- Shimoji, Michinori. 2010. Irabujimahooogen ni okeru jutsugobubun no shootenka ni tsuite [Focus marking on predicates in Irabu]. *Chikyuuken Gengokijutsu Ronshuu* 2: 115–133.
- Vance, Timothy J. 1993. Are Japanese particles clitics? *Journal of the Association of Teachers of Japanese* 27: 3–33.
- Zwicky, Arnold M. 1985. Clitics and Particles. *Language* 61: 283–305.
- Zwicky, Arnold, and Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English *n't*. *Language* 59: 502–513.

伊良部方言におけるクリティック

下 地 理 則

本研究の目的は伊良部方言のクリティック（付属語）を記述することである。通言語的にみて、クリティックという用語は音韻的に従属した単語ないしそれに類似した形式に対して用いられるが、それに準じたうえで本研究では伊良部方言の以下の形式をクリティックに認定する：格助詞、とりたて詞、副助詞、終助詞。形態統語的にみると、これらの形式は句に接続する点で接辞とは明確に異なり、一方でその出現環境の単純さ（句末）および統語規則（移動規則・削除規則）の適用状況から語とも区別される。音韻的には、ホストと同一の音韻語をなす内部付属語と、ホストが形成する音韻語の外側にある外部付属語の2種に区別できる。