

QUESTION FORMATION IN PNAR AND KHASI: THE ROLE OF CLITICS

ANISH KOSHY

Abstract

This paper is a study of the role of clitics in the process of question formation as well as in the formation of question words in two closely-related Mon-Khmer languages spoken in India, viz., Pnar and Khasi. The paper distinguishes the major types of questions in both the languages, viz., the content seeking questions and the polar questions. Within the content seeking questions, the differences and similarities in the formation of different question words are highlighted with special reference to the role of pronominal clitics. The question particles in these languages are analysed as bound forms and are classified as clitics. The process of question formation in these languages also interacts in interesting ways with the general morpho-syntax of the language which also involves the placement of agreement clitics. These languages are significant typologically and areally in the Indian context. With an SVO Basic Word Order, these two languages provide typological insights significant in the context of the Indian subcontinent, a linguistic area dominated by languages with an SOV word order. The attestation of a VSO order in the formation of polar questions, especially in Pnar, and its correlation with the SVO order follows the generally accepted typological position on VSO-SVO languages from Greenberg's studies on Universals.

Keywords: Khasi, Pnar, Mon-Khmer, typology, questions, interrogatives, VSO, SVO, clitics

Contact author:

Anish Koshy, Assistant Professor, Department of Linguistics and Phonetics,
The English and Foreign Languages University, Hyderabad, India.

Email: anish@efluniversity.ac.in

1 Introduction

Interrogatives or questions are considered one of the three fundamental types of sentences found in almost every language, the other two being declaratives and imperatives (Levinson, 1983: 40). The terms interrogative and question are often used inter-changeably in common usage. However, interrogatives across languages are known to be used for functions and purposes beyond the act of questioning (Schegloff, 1984:31; Raymond 2003: 939). This paper only deals with interrogative forms that are used in the act of questioning in the two languages under study. The two broad types of questions that are investigated, include the content-seeking questions (*wh*-questions) and the polar questions (*yes-no* questions). Typologically, languages differ in terms of the number of question words they have and also, as to how many of these question words are uniquely formed. Some also show phonological similarities across question words, with the English question words beginning with *wh*-being the most famous instance of this.

This is a descriptive analytic study of the process of question formation in two closely related Mon-Khmer languages, especially looking at the morpho-syntactic interaction of clitics with the process of question formation, based on the Basic Linguistic Theory followed in typological studies (Dixon 2012). This paper also provides important insights into the typological patterns of question formation in the two languages, in terms of the number of uniquely available or unavailable *wh*-forms, the interaction of locational, directional and case markers with *wh*-forms, and the impact of movement of constituents on the surface structure as well as on the alliterative agreement pattern.

2 Introduction to the languages

Pnar and Khasi are Mon-Khmer languages spoken in north eastern India, in the state of Meghalaya. Pnar is spoken in the Jaintia Hills District in Meghalaya. It is often called Jaintia in historical sources, the name of the community that speaks the language. Khasi is predominantly spoken in the Khasi Hills districts, which includes among other places, the state capital Shillong, and Sohra (Cherrapunjee). None of the Khasian languages are listed in Schedule VIII of the Constitution of India. According to the *Census of India 2011 report*, Khasi is spoken by 10,37,964 and Pnar by 3,19,324 people. In the sociolinguistic milieu of the region, most Pnar speakers know Khasi, as it is a language formally taught in schools and university departments. Most Khasi speakers know only rudimentary Pnar or no Pnar at all. It is also important to note that Khasi is a very generic term in the context of Meghalaya, where every variety of speech is often considered a dialect of Standard Khasi, or the rather standardized Khasi. The term *Khasi*, as used in this study, therefore, refers only to this Standardized form of the language and not to its other varieties.

3 Methodology

The data represented in this study have been mostly collected by the researcher after a field study. The data has been collected from the regions where these languages are spoken predominantly. The Pnar data was collected from Jowai and Shillong, while the Khasi data was collected from Sohra and Shillong. Most language consultants in this study are bilinguals and trilinguals, with English and Khasi as common languages. The data was collected by eliciting translations, and from narrations that were recorded and later transcribed and annotated.

4 Pnar and Khasi: a typological background

In the context of South Asia, Pnar and Khasi are important languages, as they are among the very few languages with an SVO Basic Word Order, in a linguistic landscape which is predominantly SOV. A very brief introduction to the two languages relevant to the paper are presented below.

Gender is morphological in Pnar and Khasi with all nominals carrying a gender marking. The third person pronominals are used as gender/number marking PROCLITICS on nominals as well as on other grammatical expressions, which are in agreement with these nominals (in an alliterative agreement pattern). Pronominal proclitics appear on nominals, nominal modifiers (adjectives and relative clauses), demonstratives, reciprocal pronouns, derived interrogative pronouns, derived indefinite pronouns, universal quantifiers and the numeral ‘one’. In the third person, there is a four-way distinction of masculine, feminine and epicene in the singular and that of a gender-neutral plural. The epicene singular is used with either genders to mark honorificity, diminutiveness and intimacy. This PROCLITIC is restricted to a small set of nominals, and its use is governed more by pragmatics than by syntactic/morphological factors.

Unlike Pnar, Khasi does not have a distinct set of PROCLITIC and ENCLITIC forms for each pronominal/nominal. The PROCLITIC in the verbal complex in Khasi marks the beginning of the verbal complex, parallel to the ENCLITIC closing the verb complex in Pnar. The following are the different pronominal proclitics and enclitics in Pnar and Khasi:

Proclitics (Pnar and Khasi)				Enclitics (Only in Pnar)			
3MSG	3FSG	3ESG	3PL	3MSG	3FSG	3ESG	3PL
<i>u=</i>	<i>ka=</i>	<i>i=</i>	<i>ki=</i>	<i>=u</i>	<i>=kɔ</i>	<i>=i</i>	<i>=ki</i>

A prototypical assertive sentence in Pnar with some modifiers has the following structure:

1. *u=k^hinna?* *u=wa-^hrɔŋ* *a?=u* *ja=u=sapeŋ* *u=wa=da-i?*
 3MSG=child 3MSG=RP=tall cut=3MSG ACC=3MSG=mango 3MSG=RP-PRF-ripe
 ‘The tall boy cut the ripe mango.’

The following is a prototypical assertive sentence in Khasi with some modifiers:

2. *ka=meri* *ka=daj-bam* *ja=u=sɔʔ-pieŋ* *u=ba=la-i?*
 3FSG=Mary 3FSG=PROG-eat ACC/DAT=3MSG=fruit-mango 3MSG=RP=PRF-ripe
 ‘Mary is eating a ripe mango.’

5 Question formation in Pnar and Khasi

The two types of interrogative sentences used for question formation, namely wh-questions, and polar questions involve different syntactic configurations of the pronominal clitics, and hence they are discussed separately. They also involve many other bound particles, which qualify to be clitics.

5.1 The wh-question forms in Pnar and Khasi

	Derived	Underived
‘who’	{i/u/ka/ki}=ji	-
‘whom’	ja=i=ji <i>cʰapʰaŋ</i> =i=ji	-
‘what’	{i/u/ka/ki}=ji	kamwan kamman iju
‘which’	{i/u/ka/ki}=ji {i/u/ka/ki}=wan	-
‘whose’	{i/u/ka/ki}=jɔŋ	-
‘where’	{ha/cʰa/na}=i=wan	-
‘why’	-	ile?
‘how’	-	katwan kamwan
‘when’	-	(min)nu(?) katnu

Table 1: Pnar wh-forms

	Derived	Underived
‘who’	ma=no {i/u/ka/ki}=ei	-
‘whom’	ja=no <i>sapʰaŋ=jɔŋ=nu</i>	-
‘what’	{i/u/ka/ki}=ei	aju
‘which’	{i/u/ka/ki}=no {ha/na/ja}=ka=nu	-
‘whose’	jɔŋ=nu	-
‘where’	fa=no haŋ=no na=ei	-
‘why’	-	balei
‘how’	-	kumno
‘when’	-	minno lano

Table 2: Khasi wh-forms

The tables above demonstrate that most wh-question forms in Pnar and Khasi are derived forms. In Pnar, the derivation mostly involves one of the pronominal clitics { = i/u/ka/ki} with a question particle =ji or =wan; in Khasi one of the pronominal clitics appears with the question particles =no~nu or =ei. Case markers/prepositions also interact with the question particles whenever appropriate in a very predictable and transparent manner: the GENITIVE markers =jɔŋ (Pnar) and jɔŋ= (Khasi) appear in the derivation of the possessive question word ‘whose’; the ABLATIVE, ALLATIVE and LOCATIONAL markers na=, cha=/fa=, ha=, respectively, appear in wh-forms for ‘where’ along with a question particle; the accusative marker ja= appears in the derived form for ‘whom’; the quasi-nominative marker ma= appears in the form for ‘who’ in Khasi; prepositions like *cʰapʰaŋ/sapʰaŋ* ‘about’ can host question forms that replace their nominal objects.

Many of the derived wh-forms in Pnar and Khasi have similar morpho-phonological shapes. For example, in Pnar, the question form for ‘who’, ‘what’ and ‘which’ is [PROCLITIC=ji], and in Khasi, the question form for ‘who’ and ‘what’ is [PROCLITIC=ei]. Their morpho-syntactic configurations help in their disambiguation, as will be discussed in the paper. In both Pnar and Khasi, the preferred syntactic position of the interrogative pronouns is sentence-initial, which is achieved as a result of wh-movement. Word-order is relatively free in these languages, and the wh-questions are also permitted to be in-situ.

5.2 Wh-questions in Pnar and Khasi: the morpho-syntax

In the sections that follow, we discuss the different derived wh-forms in Pnar and Khasi and the role of clitics with respect to them.

5.2.1 Questioning actors (who)

When the question forms involve the pronominal clitics, the general pattern in Pnar and Khasi with gender clitics is to copy the gender clitic from the subject noun (in Pnar questions when the head noun is unknown, the epicene *i*= is used as a default). The wh-form is followed by the RP *wa/ba* with the verb carrying no agreement markers.

3. **Pnar**
i=jì wa=pìn-jap ja=u=jɔn
 3ESG=Q RP=CAUS-die ACC=3MSG=John
 Who killed John?

4. **Khasi**
ma=no ba=pìn-jap ja=u=jɔn
 PT=Q RP=CAUS-die ACC/DAT=3MSG=John
 Who killed John? (lit: Who is the one who killed John?)

With copular sentences, we do not find the RP following the wh-form. The RP, as can be seen in the sentences above, is hosted by the verbal complex. In sentences with copular verbs, with the verbs dropped, the RP finds no host and hence is not used.

5. **Pnar**
u=jì o u=lɔk u=jɔn wa=b^ha
 3MSG=Q 3MSG 3MSG=friend 3MSG=John RP=good
 Who is John's best friend?

In Khasi, the wh-form in copular sentences has a different question particle, =ei, along with a pronominal PROCLITIC.

6. *u=ei u=n̄e*
 3MSG=Q 3MSG=PROX
 Who is he?

5.2.2 Questioning patients/objects (whom)

In both the languages, the wh-form for patients/objects is similar in shape to the one for agents with the addition of the ACC/DAT marker. The question word appears in the sentence initial position, followed by the subject nominal, if present. The verbal complex carries the subject enclitic agreement marker.

7. **Pnar**
ja=i=jì ja-c^hem=mi ha=juymane m̄intu
 ACC/DAT=3ESG=Q RECPⁱ-meet=2MSG LOC=church today
 Whom did you meet in the church today?

8. **Khasi**
ja=no p^hi=la-ja-kindu? ha=ijεŋ-manə m̄inta=ka=sŋ
 ACC/DAT= Q 2PL=PRF-RECP- LOC=home- Today
 meet worship
 Whom did you meet in the church today?

Prepositional objects answering questions like *about whom* replace the ACC/DAT marker with the preposition.

9. Pnar
 $c^h ap^h aŋ=i=ji$ $ka=meri$ $ja-klam=kɔ$
 about=3ESG=Q 3FSG=Mary RECP-talk=2FSG
 Who is Mary talking about? (Lit. *About whom is Mary talking?*)

10. Khasi
 $fap^h aŋ$ $fɔŋ=nu$ $p^h i=ja-kren$
 About GEN=Q 2PL=RECP-speak
 Whom are you talking about?

5.2.3 Questioning non-human entities and events (what)

The interrogative pronouns for ‘what’ are realized by both derived and underived forms in Pnar and Khasi. In structures with the pronominal subjects dropped, the derived interrogative form is followed by the finite verbal complex which carries the enclitic/proclitic subject agreement marker. Nominal subjects follow the wh-forms, preceding the verbal complex. It is to be noted that the derived wh-forms for ‘who’ and ‘what’ are morpho-phonologically similar, both in Pnar and in Khasi. However, the form for ‘who’ is always followed by the RP, while ‘what’ rarely comes with the RP. In Khasi, the derived interrogative form is limited to copular constructions (this is similar to the distribution of the =ei form for ‘who’ in Khasi).

11. Pnar
 $i=ji$ $jo=p^h i$
 3ESG=Q see=2PL
 What did you see?

12. Khasi
 $ka=ei$ $ka=kirteŋ$ $fɔŋ=u$
 3FSG=Q 3FSG=name GEN=3MSG
 What is his name?

5.2.4 Selection from a set (which)

In both Pnar and Khasi, the derived forms for ‘which’ contain pronominal clitics. The PROCLITIC is determined by its agreement with the gender/number of the member/members being sought from a pre-determined set, whose gender/number is already known. The questioned entity is followed by the RP hosted by the verb. The verbs carry the subject agreement enclitic/proclitic.

13. Pnar
 $ka=ji$ $ka=kɔt$ $wa=je=mi$ $c^h a=o$
 3FSG=Q 3FSG=book RP=lend=2MSG DAT=3MSG
 Which book have you lent him?

In copular constructions, with no copular verbs to host them, we do not find the RP following the wh-form.

14. Pnar
 $u=wan$ $u=paju=p^h i$
 3MSG=Q 3MSG=sibling=2SG.GENⁱⁱ
 Which one is your brother?

15. Khasi
ka=no pitar
 3FSG=Q Peter
 Which ones, Peter?

It is also possible for the wh-forms to be preceded by prepositional/case markers like the accusative, ablative, etc., if the context so requires (for example in expressions like *to which*, *from which*, etc.).

16. Pnar
na=i=wani(neiwan) i=binta ka=kɔtba? dau=pure=o
 ABL=3ESG=Q 3ESG=portion /part 3FSG=Bible FUT=read=1SG
 Which portion of the Bible should I read from?

17. Khasi
na=ka=nu ka=binta jɔŋ=ka=baibl ba ña=n=pule
 ABL=3FSG=Q 3FSG=part/portion GEN=3FSG=Bible RP 1SG=FUT=read
 Which portion of the Bible should I read? (lit. *Which is the portion of the Bible that I should read from?*)

5.2.5 Locative questions (where)

In Pnar, the derived wh-forms used for asking questions on locations ('where'), use pronominal clitics and question particles with a LOCATIONAL, ABLATIVE or ALLATIVE marker. This is a typologically well-attested pattern in many languages where deictic or motion elements get combined with interrogative words or particles (Cysouw 2004). These can be seen in English forms like *whereto*, *wherfrom*, *witherto*, etc., which are no more used as interrogative words in the language. Khasi also attests a similar pattern, except that the pronominal clitics are no more recognizable. In both the languages, speakers treat these wh-forms as non-compositional. They have become lexicalized (for example, in Khasi, it is no more possible to analyze what the particle in the middle of a form like *ha=ñ=no* is) or are in the process of becoming one (for example, the proclitic used in Pnar with these forms is invariably *i=*).

18. Pnar
heiwan(ha=i=wani) sa?=p^{hi}
 LOC=3ESG=Q stay=2PL
 Where do you stay?
19. Khasi
hañno ka=don ka=tu ka=jin̥dro
 where 3FSG=be 3FSG= DIST 3FSG= painting
 Where is that famous painting?

5.2.6 Possessive questions (whose)

Typologically, possessive questions ('whose') are derived in most languages by the addition of the possessive case marker to the question form for PERSON (*who*) (Cysouw 2004). In these two languages under study as well, the possessive questions make use of the possessive marker as part of the question word, along with a gender proclitic or a question particle. The

gender proclitic agrees with the head nominal (the possessed nominal) in Pnar. While the possessive wh-form moves to sentence-initial position in Pnar, it does not in Khasi. It remains in-situ after the possessed nominal. As expected in SVO languages, the GENITIVE in declarative sentences in both these languages, always comes following the possessed nominal.

20. Pnar

ka=jɔŋ *ka=tai* *ka=kari*
 3FSG=GEN 3FSG=DIST 3FSG=car
 Whose car is that?

21. Khasi

ka=jɪŋdro *jɔŋ=nu* *ka=ni*
 3FSG=painting GEN=Q 3FSG=PROX
 Whose painting is this?

5.3 Polar questions in Pnar and Khasi

All polar questions in Pnar have a VSO structure, with the verbs moving to the sentence-initial position. Pronominal subjects are dropped and appear as enclitics on the verb.

22. *kʷa=hi=mi* *u=kʰana* *i=ji=re=i=ji* ?
 want-EMPH=2MSG INF=tell 3ESG=Q=DRP=3ESG=Q
 Do you want to tell us anything?

23. *tɔ?=hi o* *u=laŋdɔ?*
 be=Q 3MSG 3MSG=priest
 Is he a priest?

Pnar polar questions also display VOS structures, when the object is incorporated into the verbal complex. In polar questions, it is also possible to replace the agreement enclitic on the verb with the subject nominal itself. This could suggest that the subject agreement enclitics otherwise found on the verbs, are actually traces of the subject nominal that is generated post-verbally and moves to a sentence-initial position.

24. *dau=tʰeəd-sapen=ka=meri*
 FUT=buy-mango=3FSG=mary
 Will Mary buy mangoes?

In polar questions with modal verbs, some modals take the agreement enclitic, while others require the main verb to take the enclitic. Pronominal subjects are dropped. In Pnar, modal verbs are followed by lexical verbs marked with a non-finite marker. Incorporated objects result in VOS structures.

25. *je=me* *u=e* *u=kʰulɔm* *ja=ŋa*
 MOD:ABIL=2MSG INF=give 3MSG=pen DAT=1SG
 Can you give me a pen?

26. *hoi=hi* *u=di?-duma=ɔ* *hεini*
 MOD:PERM=Q INF=drink-cigarette=1SG here

Can I smoke here?

The formation of polar questions is a very distinguishing syntactic feature between Pnar and Khasi. Khasi employs a question word *hatɔ?* in the sentence-initial position for all types of polar questions. The question word is followed by the assertive sentence without any interaction with pronominal proclitics. The rising intonation conveys that it is a question.

27. *ʃɔn, hatɔ? ɳi=n=leit da=ka=kali ʃɔɳ=u*
 John Q 1PL=FUT=go INS=3FSG=car GEN=3MSG
 Are we going by your car John?

28. *hatɔ u=ne u=ʃu u=brijəu u=ba ɳa=la-ja-kindɔ? menʃim-snem?*
 ? Q 3MSG= PROX 3MSG= Same 3MSG= 3MSG= 1SG= PRF-
 person RP RECP-meet

'Is this the same person whom I met last year?'

6 Conclusion

Pnar and Khasi are two important languages within the Mon-Khmer branch of Austroasiatic. While a lot of studies have been undertaken on Khasi (cf: Rabel 1961; Nagaraja 1985), there is very little research available on Pnar. This comparative study of the interaction of clitics with the process of question formation, as well as the role of clitics in the formation of question words in these two languages, tries to cover some of this gap. There are many questions that still remain unanswered. Why do questions with lexical verbs need a relativized clause to follow them? Why do questions with copular verbs not need these relativized clauses? These issues need further investigations.

The study has also been able to demonstrate significant differences between the two varieties of Mon-Khmer speech, spoken in close proximity. The important morpho-syntactic differences highlighted in the study are important in the larger study of the Khasian branch of languages, as well as contributes to the typological understanding of questions and question formation. The role of clitics as highlighted in the paper, is hoped to also serve in developing the morpho-syntactic understanding of clitics in the literature, which are mostly based on European languages. It is also hoped that the study will enable a more holistic understanding of the two languages within the context of the Mon-Khmer subfamily.

List of abbreviations used

-	AFFIX BOUNDARY	INF	INFINITIVE
=	CLITIC BOUNDARY	INS	INSTRUMENTAL
1	FIRST PERSON	LOC	LOCATIVE
2	SECOND PERSON	M	MASCULINE
3	THIRD PERSON	MOD:ABIL	ABILITY MODAL
ABL	ABLATIVE	MOD:PERM	PERMISSION MODAL

ACC	ACCUSATIVE	PL	PLURAL
CAUS	CAUSATIVE	PRF	PERFECTIVE
DAT	DATIVE	PROG	PROGRESSIVE
DIST	DISTAL	PROX	PROXIMATE
DRP	DISCONTINUOUS REDUPLICATION PARTICLE	P	PARTICLE
E	EPICENE	Q	QUESTION
EMPH	EMPHATIC	RECP	RECIPROCAL
F	FEMININE	RP	RELATIVIZING PARTICLE
FUT	FUTURE	SG	SINGULAR
GEN	GENITIVE		

References

- Census of India. (2011) Series 1: Language India and states. Registrar General and Census Commissioner, India.
- Cysouw, M. (2004) Interrogative words: An exercise in lexical typology. In *Presentation presented at the Bantu grammar: description and theory workshop*, (Vol. 13).
- Dixon, R. M. W. (2012) *Basic Linguistic Theory*, Vol. 3. Oxford: OUP.
- Emanuel A. S. (1984) On some questions and ambiguities in conversation. In J. Maxwell Atkinson and John Heritage (eds) *Structures of Social Action: Studies in Conversation Analysis* (pp. 28-52). Cambridge: Cambridge University Press.
- Levinson, S.C. (1983) *Pragmatics*. Cambridge: Cambridge University Press.
- Nagaraja, K.S. (1985) *Khasi: A Descriptive Analysis*. Pune: Deccan College Post-Graduate and Research Institute.
- Rabel, Lili. (1961) *Khasi, A Language of Assam*. Baton Rouge: Louisiana State University Press.
- Raymond, G. (2003) Grammar and social organization: Yes/no interrogatives and the structure of responding. *American sociological review* 68 (6), 939-967.