Chapter 10

Preposed NPs in Seediq

Naomi Tsukida
Aichi Prefectural University

Seediq is an Austronesian language spoken in northeastern Taiwan. Its word order is VXS in general (where X stands for adjuncts or arguments other than the subject), but an NP that has some semantic relation to the matrix clause can precede the matrix clause. This NP is followed by the particle *'u* and a non-final pause. I will call such NPs preposed NPs. This paper will investigate the nature of these preposed NPs. What is their semantics and function? What type of NPs can be preposed? What are their anaphoricity and topic-persistence properties? Preposed NPs are often coreferential with the matrix subject, which may or may not be overt. When a preposed NP is coreferential with the main clause subject, and the matrix subject is not overt, then the seeming word order is SVX. When the preposed NP is coreferential with the matrix subject, and that subject is overt, then the seeming word order is SVXS. How do SVX and SVXS differ? What is the function of this double reference in SVXS? In addition, how do these word orders differ from simple VXS order?

1 Introduction

Seediq is an Austronesian language, spoken in northeastern Taiwan and belonging to the Atayalic subgroup. There are three Seediq dialects: Teruku, Tekedaya, and Tuuda. The research reported here is based on the Teruku dialect, which is mainly spoken in Hualien County. The population of the Teruku subgroup is about 30,000, but the younger generations do not speak the language.

In what follows §1.1 explains Seediq word order. Subject and voice are covered in §1.2, non-subject arguments in §1.3, and the *'u* particle in §1.4.

1.1 Word order

The basic constituent order in Seediq is VXS (X stands for adjuncts or arguments other than the subject). In example (1), *kerut*[1] '<AV>cut' is the V, *bunga* 'sweet potato' is a non-subject argument, and *payi* 'old woman' is the subject. *Ka* is multi-functional. Its function here is to mark the subject.

[1]The phoneme inventory of Teruku Seediq is as follows: p, t, k, q, ', b, d, s, x, h, g (voiced velar fricative), c, l (voiced lateral fricative), r, m, n, ng (velar nasal), w, y, a, i, u, and e (schwa).

Naomi Tsukida. 2018. Preposed NPs in Seediq. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Perspectives on information structure in Austronesian languages*, 313–344. Berlin: Language Science Press. DOI:10.5281/zenodo.1402553

(1) AV
Kerut bunga ka payi.
<AV>cut sweet.potato SBJ old.woman
'The old woman cuts sweet potato.'

In NPs, modifiers other than numbers typically follow the head. See §3.1 for details.

Seediq has two sets of clitic pronouns, NOMINATIVE and GENITIVE, and two sets of independent pronouns, NEUTRAL and OBLIQUE. While all the pronouns have a GENITIVE clitic, the NOMINATIVE clitic is limited to the first and second persons. Clitic pronouns are second-position clitics, following the first element of the predicate. Some of the examples are *=ku* in example (2) and *=na* in examples (6), (7), and (8).

(2) Clitic pronoun
Kerut=ku bunga (ka yaku).
<AV>cut=1SG.NOM sweet.potato SBJ 1SG
'I cut sweet potato.'

One can omit adjuncts and arguments including the subject if they are recoverable from the context. Clitic pronouns are enough to indicate arguments, so when there is a nominative clitic in the sentence, it does not need to be referred to by an independent pronoun.

In addition to the basic VXS order, Seediq can have an NP preceding the matrix clause, followed by the particle *'u* and a non-final pause. In example (3), *payi* 'old woman' is preposed and followed by the *'u* particle and a non-final pause. The matrix subject *ka payi* may be omitted, because, even if it is omitted, it is identifiable as coreferential with the preposed NP.

(3) Preposed NP
Payi 'u, kerut bunga (ka hiya).
old.woman PRT <AV>cut sweet.potato SBJ 3SG
'The old woman, she cuts sweet potato.'

As a result, SVX order (more precisely, *S 'u, V X* (*ka S*)) is realized, though I do not regard the preposed NP as a subject. This paper will investigate the nature of such preposed NPs.

1.2 Subject and voice

The subject appears in the Neutral case, preceded by the particle *ka*, in clause-final position. *Ka* is multi-functional; some of its functions are subject-marker, linker, and complementizer. Its usage as a linker is yet to be investigated.

In Seediq, thematic roles fall into three groups, according to the verb form they trigger when they are chosen as the subject. The first group, the A group, includes AGENT, THEME, and EXPERIENCER; these trigger AV forms, as in (1) in §1.1. The second group, the G group, includes PATIENT, GOAL, LOCATION, and RECIPIENT; these trigger GV forms, as in (4). The third group, the C group, includes CONVEYED THEME, INSTRUMENT, and

Beneficiary; these trigger CV forms, as in (5). In (4) and (5), the patient *bunga* 'sweet potato' and the instrument *yayu* 'knife' is the subject, respectively, and the Agent *payi* 'old woman' is not the subject anymore.

(4) GV
Kerut-un payi ka bunga.
cut-GV1 old.woman SBJ sweet.potato
'The/An old woman will cut the sweet potato.'

(5) CV
Se-kerut bunga payi ka yayu.
CV-cut sweet.potato old.woman SBJ knife
'The/An old woman cut the/a sweet potato with the knife.'

Patient and location are grouped together; forms suffixed with *-un* are GV1, and those with *-an* are GV2 forms. Corresponding forms in other languages are often regarded as a patient voice and a location voice, respectively (Huang & Hsinsheng 2016: 78, for example), but in Teruku-Seediq, their usage is not so straightforward. Even when the thematic role of the subject is location, if one wants to express a future event, the *-un* form is used, for example. Moreover, even when the thematic role of the subject is patient, the *-an* form may be used if one wants to express a progressive or habitual event. See Tsukida (2012) for more details.

When the subject is a pronoun, it triggers a nominative clitic pronoun after the first element of the predicate (see example (2)).

There are cases where the thematic role of a *ka*-marked NP and the verb form do not correspond. For example, the verb is in AV form but the thematic role of the *ka*-marked NP is patient. In such cases the *ka*-marked NP cannot be regarded as the subject. Such cases exist, though they are rare.

In addition to triggering the nominative clitic pronoun, subjects show several distinctive morpho-syntactic properties (see Tsukida 2009 for details).

1.3 Non-subject arguments

An NP belonging to the A group (the group of NPs with thematic roles that trigger AV) is realized by the genitive clitic if it is a pronoun, when it is not the subject. The A arguments in (4) and (5) are not subjects. If one replaces them with pronouns, they would be as (6) and (7), respectively.

(6) GV, with pronominal A
Kerut-un=na ka bunga.
cut-GV1=3SG.GEN SBJ sweet.potato
'He/She will cut the sweet potato.'

(7) CV, with pronominal A
Se-kerut=na bunga ka yayu.
CV-cut=3SG.GEN sweet.potato SBJ knife
'The/An old woman cut the/a sweet potato with the knife.'

As for nouns, they do not have a distinct genitive form; the genitive form is the same as the neutral form. We can regard *payi* 'old woman' in (4) and (5) as a genitive form, though it is formally the same as a neutral form.

When it is not the subject, an NP belonging to the G or C group (the group of NPs with thematic roles that trigger GV or CV) is realized as OBLIQUE. The oblique form, however, is not distinct from the neutral form, except for pronouns and nouns with high animacy. For pronouns and NPs with high animacy, the oblique form usually involves the suffix *-an*. The occurrences of *bunga* 'sweet potato' in (1), (2), (3), (5) and (7) are regarded as oblique forms, though they are homophonous with neutral forms. *Sediq-an* 'person-OBL' in (8) is an example of a distinct oblique form.

(8) CV
Se-kerut=na sediq-an ka yayu niyi.
CV-cut=3SG.GEN person-OBL SBJ knife PROX
'He/She cuts the/a person with this knife.'

1.4 The *'u* particle

As mentioned in §1.1, an NP may appear preceding a matrix clause, followed by the particle *'u* and a non-final pause. The particle may be *ga*, *de'u*, or *dega*, in addition to *'u*. *Ga* is interchangeable with *'u*. I will use only *'u* in the remainder of the paper. *De'u* and *dega* are interchangeable, as they are derived from *'u* and *ga* by affixation of *de-*. *De'u* and *dega* are used differently from *'u* and *ga*. I will not treat the use of *de'u* and *dega* in this paper.

What can precede *'u* is not only an NP but also a clause. In example (9), two clauses are connected by *'u*.

(9) Clause-A 'u, clause-B.
M-iyah=su hini 'u, me-qaras=ku.
AV-come=2SG.NOM here PRT AV-be.glad=1SG.NOM
'If you come, I will be glad.'

'u is multifunctional, and the two clauses connected by it may have several kinds of relationships. In this sentence, the preposed clause is a conditional for the event denoted by the matrix clause. One can add an appropriate adverb to express the relationship overtly. In example (10), for example, *nasi* 'if' is added to express the conditional meaning overtly.

(10) With *nasi* 'if'
Nasi=su m-iyah hini 'u, me-qaras=ku.
if=2SG.NOM AV-come here PRT AV-be.glad=1SG.NOM
'If you come here, I will be glad.'

1.5 Texts

In the next section, I will investigate the nature of the preposed NPs mentioned in §1.1 from the point of view of semantics/function, NP types, anaphoricity, and topic persistence, analyzing Seediq texts.

For my analysis, I used parts of *Sufferings of the Teruku church* written by Yudaw Pisaw (Tien Shin-de). In these texts, matrix clause subjects occur five times more frequently than preposed NPs. I started checking from the beginning and when I reached 3728 words, 262 matrix clause subjects and 55 preposed NPs had been found. I therefore stopped counting matrix clause subjects but continued counting preposed NPs, up to 7315 words, and found 38 more instances of preposed NPs. In total, I have 93 instances of preposed NPs.

Table 1: Matrix subjects and preposed NPs in texts

Words	Matrix subject	Preposed NP
1–3728	262	55
3729–7315	– (stopped counting)	38
Total	262	93

2 Semantics/function of preposed NPs

It is not the case that one can freely choose an NP to precede the matrix clause. Preposing is usually restricted to the following NPs:

- An NP that is coreferential with the matrix subject (preposed subject)
- An NP that is coreferential with the A-argument in the matrix clause (preposed A)
- An NP that expresses the place of existence in an existential construction, or the possessor in a possessive construction, or a possessor that is left-dislocated from the matrix subject (preposed possessor, see §2.3 for details)
- An NP that expresses time (preposed time)

- Choices in alternative questions are preposed even when they are not subjects, A-arguments, times, places, or possessors (preposed alternatives)

Some preposed NPs apparently do not fit the above criteria. They are not coreferential with a subject, an A argument, or a possessor in the matrix clause; nor are they frames or alternatives. These will be discussed in §2.6.

Among 93 instances of preposed NPs, 66 instances have a preposed NP that is coreferential with the subject of the clause (preposed subject); in four instances it is coreferential with an A-argument (preposed A), in one instance it is a possessor (preposed possessor), and in 16 instances it is time (preposed time). There were six instances where it was difficult to judge the exact function of the preposed NP. These 93 instances are summarized in Table 2.

Table 2: Semantics/function of preposed NP

	Token	%
Preposed Subject	66	70%
Preposed A	4	6%
Preposed Possessor	1	1%
Preposed Time	16	17%
Preposed alternative	0	0%
None of the above	6	6%
Total	93	100%

I will explain each of the above cases in turn.

2.1 Preposed subjects

One can prepose an NP that is coreferential with the matrix subject, as in (3) and (11). I will call such cases PREPOSED SUBJECTS, though preposed NPs are not subjects, actually.

(11) Preposed subject
Niyi$_{i}$ *'u, 'adi 'utux=ta* (*ka kiya*$_{i}$).
PROX PRT NEG God=1PLIN.GEN SBJ it
'As for this, it is not our God.'

One can omit the matrix subject, as shown in (11).

When an NP that is coreferential with the subject appears in pre-clausal position, the subject in the regular clause-final position is often omitted. In the text, 13 of 66 preposed subjects had overt subjects, and 53 instances had covert subjects. I will investigate in §3.3, §4.3, and §5.3 whether there is any difference between these two cases.

With a verbal predicate, "subject" means the NP that triggers the verb form, appearing in clause-final position preceded by the subject particle *ka*. *Payi* 'old woman' in (1), *bunga*

'sweet potato' in (4), and *yayu* 'knife' in (5) are examples of subjects. For sentence (1), one can put *payi* 'old woman' in front of the matrix clause, as in (12), but one cannot do the same for *bunga* 'sweet potato,' as in (13).

(12) One can prepose a subject argument
Payi 'u, kerut bunga (ka hiya).
old.woman PRT <AV>cut sweet.potato SBJ 3SG
'The old woman, she cuts sweet potato.'

(13) One cannot prepose a non-subject argument
**Bunga 'u, kerut ka payi.*
sweet.potato PRT <AV>cut SBJ old.woman
'The old woman cuts sweet potato.'

For sentence (4), one can prepose *bunga* 'sweet potato', as in (14).

(14) GV
Bunga 'u kerut-un payi (ka bunga).
sweet.potato PRT cut-GV1 old.woman SBJ sweet.potato
'The/An old woman will cut the sweet potato.'

A preposed subject is not an actual subject, as I have indicated, but only coreferential with the matrix subject. It seems to share some of the subject properties mentioned in §1.2, however. A nominative clitic that corresponds to the preposed NP appears in the main clause, for example.

(15) One can prepose a subject argument
Yaku 'u, kerut=ku bunga (ka yaku).
1SG PRT <AV>cut=1SG.NOM sweet.potato SBJ 1SG
'As for me, I cut sweet potato.'

We cannot tell, however, whether it is the preposed NP *yaku* or the subject of the matrix clause that triggers the clitic.

When the preposed NP is coreferential with the matrix subject, the matrix subject may either appear again or be omitted. For 13 of 66 preposed subject, matrix clause subject is overt and for 53 it was not overt. When it is overt, it is realized either by *hiya*, the 3rd person singular pronoun, by *dehiya*, the 3rd person plural pronoun, or by *kiya* 'it, so'. In example (16), *Karaw Wacih* (proper name) appears in preposed position and at the end of the matrix clause, it is referred to again by a pronoun *ka hiya*.

(16) Preposed subject with overt matrix subject
Si'ida ka Karaw Watih niyi 'u adi ka balay=bi senehiyi Yisu Kiristu
then LNK Karaw Wacih PROX PRT NEG LNK really=really AV.believe Jesus Christ
ka hiya niyana.
SBJ 3SG yet
'At that time Karaw Wacih, he did not really believed in Gospel yet.'

I will call this pattern *S 'u, VX ka hiya*. In this denotation, *hiya* represents *hiya*, *dehiya* and *kiya*. The pattern where the matrix subject is not overt, on the other hand, will be labeled *S 'u, VX*.

2.2 Preposed A

Here by "A" I mean the group of NPs which are not the subject but whose thematic roles trigger AV when they become the subject. A pronoun is in the genitive case when it appears as non-subject, as in (6) and (7).

An A argument can be preposed to the matrix clause. In a sentence like (5), for example, one can prepose *payi* 'old woman', as in (17), because its thematic role is agent. The genitive clitic pronoun *=na*, which is coreferential with *payi* appears in the matrix clause.

(17) CV
Payi 'u, se-kerut=na bunga ka yayu.
old.woman PRT CV-cut=3SG.GEN sweet.potato SBJ knife
'As for the old woman, she will cut the/a sweet potato with the knife.'

There are four such instances among the 95 preposed NPs.

Preposing A is not allowed if there is no X (non-subject argument). In (18) there is only the A argument *Kumu* (a person's name) as a non-subject argument. From this sentence, one cannot prepose A, as shown in (19).

(18) A as non-subject argument
B<en>arig Kumu ka patas niyi.
<CV.PRF>buy Kumu SBJ book PROX
'Kumu bought this book.'

(19) One cannot prepose A
**Kumu 'u, b<en>arig=na ka patas niyi.*
kumu PRT <CV.PRF>buy=3SG.GEN SBJ book PROX
'As for Kumu, she bought this book./As for Kumu, this book was bought by her.'

A time expression is not enough to enable preposing of A. Even if *sehiga* 'yesterday' is X, one cannot prepose A, as shown in (20).

(20) One cannot prepose A
**Kumu 'u, b<en>arig=na sehiga ka patas niyi.*
Kumu PRT <CV.PRF>buy=3SG.GEN yesterday SBJ book this
'As for Kumu, she bought this book yesterday.'

2.3 Preposed possessor

One of the morphosyntactic properties of subjects is that they license possessor left-dislocation. From a subject NP consisting of a noun and a possessor, one can move out

the possessor and put it in front of the clause. Examples (21) and (22) exemplify this. In (21) *Masaw* (the name of a male person) is part of the NP *tederuy Masaw* 'Masaw's car', which is the subject of the sentence. In (22), on the other hand, *Masaw* is preposed, and in the subject a genitive clitic pronoun is present which is coreferential with the preposed *Masaw*.

(21) *Me-gerung ka tederuy Masaw.*
AV-be.broken SBJ car Masaw
'Masaw's car is broken.'

(22) *Masaw 'u, me-gerung ka tederuy=na.*
Masaw PRT AV-be.broken SBJ car=3SG.GEN
'As for Masaw, his car is broken.'

There is one such instance among the 93 preposed NPs.

2.4 Preposed time

Preposed time sets the temporal frame in which the following expression should be interpreted. An example is (23).

(23) Preposed time
Ya'asa diyan 'u, m-iyah rigaw sbu ka 'asu sekiya
because day:time PRT AV-come <AV>hang:about <AV>hit SBJ ship sky
'Amirika heki.
America REASON
'Because in the daytime, American airplanes come to make an air raid.'

There are 16 such instances among the 93 preposed NPs.

2.5 Preposed alternatives

For alternative questions, alternatives are preposed, as in (24).

(24) Preposed Alternatives
Deha niyi 'u, 'ima ka sewayi=su?
two PROX PRT who SBJ younger:sibling=2SG.GEN
'Between these two, who is your younger sibling?'

There was no such example in the text I used in this investigation.

2.6 None of the above

Some preposed NPs apparently do not belong to any of the classifications listed above. Such NPs rarely become subjects, but examples do exist. In (25), *saw niyi* 'such things' is a patient, not A, of the verb *me-kela* 'to know', notionally. It is not a subject, either, because the predicate verb is AV, not GV, which would have a patient as its subject. Nor is it a possessor, time or expression of alternatives.

(25) None of the above
Saw niyi 'u, dima me-kela ka nihung.
like PROX PRT already AV-know SBJ Japan
'As for such things, Japan already knew.'

There were six such instances in the text I analyzed.

2.7 Discussion

Let us examine here whether these preposed NPs are "topics" of some sort.

Preposed subjects, preposed As, and preposed possessors, shown in §2.1, §2.2 and §2.3, seem to serve as aboutness topics (Krifka 2007: 40–41). They indicate what the information denoted by the matrix clause (=comment) is about. Examples (12) and (17), for example, are about *payi* 'old woman'. The matrix clause tells the information about the preposed NP.

Preposed time and preposed alternative, on the other hand, seem to serve as frame setters. The frame setter indicates that the information actually provided is restricted to the particular dimension specified (Krifka 2007: 47). In example (23), the information provided by the matrix clause 'American airplanes come to make an air raid' is restricted to the particular dimension specified by the preposed NP *diyan* 'daytime', for example.

For those elements that are neither preposed subjects, preposed As, preposed possessors, preposed time or preposed alternatives, still more research is necessary, but as for example (25), the preposed NP, *saw niyi* 'such things' seems to serve to set frame for the information provided by the matrix clause, 'Japan already knew'.

3 NP types

In this section, I will investigate the composition of preposed NPs.

3.1 Preliminaries

In Seediq NPs, a noun (a common noun or a proper noun) or a pronoun usually functions as the head, but demonstratives and VPs can also function as referential expressions, without any affix or particle. Seediq is quite flexible, in the sense of van Lier & Rijkhoff (2013).

(26) A noun as a referential expression
Qita=ku kuyuh ka yaku.
<AV>see=1SG.NOM woman SBJ 1SG
'I saw a/the woman.'

(27) A demonstrative as a referential expression
Me-gerung ka gaga.
AV-be:broken SBJ DIST
'That is broken.'

VPs also are used as referential expressions without any additional affixation or particle. AV forms mean 'one who does the action denoted by the verb'/'one who bears the state denoted by the verb'; GV forms mean 'the object to which the action denoted by the verb is done'/'the place where the action denoted by the verb takes place'/'the recipient who recieves something'/'the goal toward which the action denoted by the verb aims'/ etc.; CV forms mean 'the instrument by which the action is done'/'the beneficiary for whom the action is done'/'the object that is transferred'/etc." A CV-perfect form may mean 'what has been done as the action denoted by the verb' (see (48)). *Mpe-tegesa kari* 'the one who teaches language' in (28) is an example of an AV form used as a referential expression.

(28) A VP as a referential expression
Qita=ku mpe-tegesa kari ka yaku.
<AV>see=1SG.NOM AV.FUT-teach language SBJ 1SG
'I saw the language teacher.'

When the head is a noun, it can be modified by demonstratives, numbers, genitive pronouns, another noun, or relative clauses. Modifiers, except for numbers, typically follow the head.

(29) A demonstrative modifying a noun
Qita=ku kuyuh gaga ka yaku.
<AV>see=1SG.NOM woman DIST SBJ 1SG
'I saw that woman.'

(30) A number modifying a noun
Me-gerung ka teru tederuy.
AV-be:broken SBJ three car
'The three cars are broken.'

(31) A genitive pronoun modifying a noun
Qita=ku kuyuh=na ka yaku.
<AV>see=1SG.NOM wife=3SG.GEN SBJ 1SG
'I saw his wife.'

(32) Another noun modifying a noun
Me-gerung ka sapah qehuni.
AV-be:broken SBJ house wood
'The wooden house is broken.'

(33) A relative clause modifying a noun
Qita=ku kuyuh ***mpe-tegesa*** *kari ka yaku.*
<AV>see=1SG.NOM woman AV.FUT-teach language SBJ 1SG
'I saw the woman who will teach language.'

I classified NPs into the following types, and investigated how they appear in preposed position. I also investigated their distribution as subjects and made comparisons.

- Personal pronoun (PP)
- Demonstrative (D)
- Proper noun + *niyi* (this) (PD)
- Common noun + *niyi* (this) (CD)
- Bare proper noun (BPN)
- Common noun + Genitive (CG)
- Common noun + Relative clause/Attributive (CR)
- Common noun (CN)
- *Saw* nominal (*Saw*) 'Things like …'
- VP

This classification and ordering are based in part on Gundel et al.'s (1993) classification which claims that the form of a given NP reflects the givenness of the NP and show the Givenness Hierarchy (Gundel et al. 1993: 275).

in focus	>	activated	>	familiar	>	uniquely identifiable	>	referential	>	type identifiable
{*it*}		{*that*, *this*, *this* N}		{*that* N}		{*the* N}		{indefinite *this* N}		{*a* N}

Figure 1: Givenness hierarchy (Gundel et al. 1993: 275)

In Seediq, too, a personal pronoun (PP) may reflect *in focus* status, which is at the left-most position in the hierarchy, demonstrative (D) may reflect *activated* status, common noun + demonstrative (CD) may reflect *activated* or *referential* status, and so on. As

Seediq does not have a definite article or an indefinite article, it is not easy to identify each NP as uniquely identifiable or type identifiable

At this stage, it is an open question whether and how the different composition in Seediq NPs reflects givenness. This is still under investigation, and the classification above is only a preliminary step. Though still to be examined more thoroughly, I tried ordering them so that it would correspond, at least partially, to the givenness hierarchy by Gundel et al. (1993).

I will explain my classification and ordering below.

First is an NP consisting of a personal pronoun. In Gundel et al. (1993), this is an indication of *in focus* status. An example of a personal pronoun preposed to a matrix clause:

(34) Personal pronoun (PP)
Hiya 'u, ya'a=bi teru 'idas ka ne-niq-un=na hiya ni
3SG PRT UNCERTN=really three month SBJ FUT-live-GV1=3SG.GEN there and
'Then she, it was about three months that she would live there, and ….'

Next is an NP consisting of a demonstrative. In Gundel et al. (1993), this is an indication of *activated* status. An example of a demonstrative preposed to a matrix clause:

(35) Demonstrative (D)
Niyi 'u, 'adi 'utux=ta ka niyi.
PROX PRT NEG God=1PI.GEN SBJ PROX
'This, this is not our God.'

I also treated the time expression *si'ida* 'then, at that time' as demonstrative.

(36) Demonstrative (D)
Si'ida 'u, me-seseli ngangut sapah 'Umih Yadu.
then PRT\ AV-gather outside house PN PN
'At that time, they gathered outside of the house of Umih Yadu.'

Next is an NP consisting of a proper name and a demonstrative (PD). Gundel et al. (1993) do not include such a form. An example of a proper noun modified by a demonstrative appearing in front of the matrix clause:

(37) Proper noun + *niyi* (PD)
Ya'asa Tiwang niyi 'u, m-en-niq degiyaq qawgan.
because Ciwang PROX PRT AV-PRF-live mountain Qawgan
'Because this Ciwang, she was a Teruku, and she had lived at Mt. Qawgan.'

Proper nouns modified by demonstratives seem to be *activated* or *familiar*.

Next is an NP consisting of a common noun and a demonstrative. Gundel et al. (1993) point out the functional difference between *this* and *that* from the point of view of information structure. According to them, *this N* is an indication of *activated* status, but *that*

N indicates *familiarity*. I examined the corpus to determine whether the corresponding expressions in Seediq, *N niyi* (*this N*) and *N gaga* (*that N*), show a similar difference, but I found only instances of *N niyi*. There are no instances of *N gaga* in the text, but *N gaga* is often found in elicitation. An example of a preposed CD (common noun modified by a demonstrative):

(38) Common noun + *niyi* (CD)
Budi pe-patas=mu niyi 'u, biq-i=ku haya
bamboo FUT-CV.write=1SG.GEN PROX PRT give-GV.NFIN=1SG.NOM BEN
kuyuh=mu rubiq wilang ha.
wife=1SG.GEN PN PN GENTLE
'As for this fountain pen, please give it to my wife Rubiq Wilang.'

Common nouns modified by demonstratives (CDs) are supposed to be *activated*, *familiar* or *uniquely identifiable*.

Next is an NP consisting of a proper name. An example:

(39) Bare proper noun (BPN)
Kiya de'u, nihung 'u, hengkawas 19,14 m-adas hebaraw=bi merata ni kensat
so PRT Japan PRT year 1914 AV-bring many=really soldier and police
m-iyah keremux dexegal Teruku.
AV-come <AV>invade land Teruku
'Then, Japan, in 1914, brought soldiers and police and came to invade the Truku territory.'

Things or persons denoted by proper names are at least *uniquely identifiable*.

Next is an NP consisting of a common noun and a genitive pronoun. The genitive expresses a possessor, as illustrated in (40).

(40) Common noun + Genitive pronoun (CG)
Bubu=na 'u, wada m-arig bawa da.
mother=3SG.GEN PRT be:gone AV-buy bun NS
'His/Her mother went to buy a bun.'

Common nouns modified by genitive pronouns (CGs) are supposed to be *uniquely identifiable* or *referential*.

Next is an NP consisting of a common noun and a relative clause or attributive expression. In Seediq, a noun may be modified by another noun, and the relationship between them is possessor-possessed, material-thing (e.g., paper napkin), purpose-thing (e.g., traveling shoes), whole-part (e.g., face tattoo), product-producer (e.g., his book, meaning 'the book he wrote'), etc. I collectively call those meanings that are expressed by the modifiers illustrated "attributive". "Relative clause" is actually a VP modifying a noun (see (42)). This type, therefore, is a class of NPs consisting of a noun and another noun or a VP modifying it. Examples are:

(41) Common noun + Other noun (CR)
Kana ka mensewayi de-senehiyi 'uri 'u, tetegeli'ing m-usa me-seseli
all LNK brother PL-believer also PRT AV.hide AV-go AV-gather
tehemuku.
AV.worship
'The brothers and sisters of believers went together hiding to worship as well.'

(42) Common noun + VP (CR)
Kensat nihung m-eniq Sekadang 'u, Matsudo-sang ni Motoyoshi-singsi 2
police Japan AV-live Sekadang PRT Matsudo-Mr. and Motoyoshi-teacher two
hiyi.
body
'The Japanese police who were at Skadang, they were Mr. Matsudo and Teacher Motoyoshi, two people.'

Common nouns modified by another noun or VP (CRs) are also supposed to be *uniquely identifiable*, *referential* or type *identifiable*.

Next is an NP consisting of a common noun alone, as illustrated in (43):

(43) Common noun (CN)
Ke'man 'u, pengkeku'ung sengkekingal m-usa 'ayug daya 'alang=deha
night PRT in:the:darkness one:by:one AV-go stream upwards village=3PL.GEN
ka se'diq senehiyi Kiristu.
SBJ person AV-believe Christ
'At night, people who believed in Christ went to the upper stream of their village, in the darkness, one by one.'

Bare common nouns may be *uniquely identifiable*, *referential*, or *type identifiable*. Seediq does not have any definite article or indefinite article, so it is often hard to tell.

Next is *saw + noun*. An NP is preceded by *saw*, a preposition meaning 'like …'; *saw NP* means 'things like an NP' or 'things concerning NP.' Sometimes *ka* appears between *saw* and *NP*. An example:

(44) *Saw* nominals
Saw ka qaya samat 'u, des-un=deha be'nux, seberig-an
like LNK thing wild:animal PRT bring-GV1=3PL.GEN plain:land sell-GV2
kelemukan.
Taiwanese
'As for things concerning wild animals (animal skins, bones, horns, and the like), they brought them to the plains and sold them to Taiwanese.'

Saw + noun would correspond to *type identifiable* in givenness hierarchy by Gundel et al. (1993).

The last is VP. A VP can function as a referential expression, as in example (28). This type of expression may appear as the subject or as a non-subject argument, but rarely as a preposed element. (45) is another example of a VP functioning as a subject phrase. In this example, a VP *n-arig=mu sehiga* '<CV.PRF>buy=1SG.GEN yesterday' means 'what I bought yesterday.'

(45) VP as matrix subject
Me-gerung ka n-arig=mu sehiga.
AV-be:broken SBJ CV.PRF-buy=1SG.GEN yesterday
'What I bought yesterday is broken.'

Such VPs used as referential expression may be *uniquely identifiable*, *referential* or *type identifiable*, depending on the adjuncts contained in them.

3.2 Results

I investigated what type of NP appears for each type of preposed NP classified in §2. The results are shown in Table 3.

Table 3: Types of preposed NPs, in terms of coreference

	Preposed				Others	Sum
	Subject	Time	A	Po		
PP	2	0	0	0	0	2
D	6	10	0	0	2	18
PD	12	0	0	0	0	12
CD	4	0	0	0	0	4
BPN	20	0	2	0	0	22
CG	6	0	1	1	0	8
CR	4	0	0	0	1	5
CN	4	6	1	0	0	11
Saw+ noun	7	0	0	0	3	12
VP	1	0	0	0	0	1
Interrogative	0	0	0	0	0	0
Total	66	16	4	1	6	

There are several things I can point out from the results above.

Personal pronouns rarely appear preposed. There are only two instances of them.

For preposed subjects, person names, accompanied by a demonstrative (PD) or not (BPN), appear more often than the other types of NPs.

For preposed time, there are 10 instances of *si'ida* 'then, at that time.' These are classified as demonstrative. The other six instances are bare common nouns (CN).

Among those six instances that were not S, A, Po, or Time, three are of the type *saw* + noun 'things concerning *noun*.' It seems easier for speakers to ignore the semantic relationship between the preposed NP and the voice of the predicate verb of the matrix clause when the preposed NP is *saw* + noun. Let us examine here the three cases more precisely. One is example (25). The other two are (46) and (47).

(46) Saw + noun
Saw niyi 'u, berah 'ini 'iyah dexegal Taywan hini ka Nihung
like PROX PRT before NEG AV.NFIN.come land Taiwan here SBJ Japan
han.
temporarily
'As for such things, it was before Japan came here to the land of Taiwan.'

(47) Saw + noun
Saw niyi 'u, 'adi=nami tegesa ka yami kensat Nihung 'uri pini,
like PROX PRT NEG=1PLEX <AV>teach SBJ 1PLEX police Japan also ANGER
mesa.
HEARSAY
'As for such things, we Japanese police do not teach, they say.'

In these examples, *saw niyi* 'such things' seems to indicate rather abstract situations, not concrete entities. Such lack of concreteness seems to lead to inconsistency of the voice form of the matrix verb.

Below, I will investigate whether the situation differs for preposed subjects if the matrix subject is overt or not. I will also compare the preposed subject situation with the situation of the matrix subject.

3.3 Comparison

Among sentences with preposed subjects, some have an overt matrix subject, and some do not. When the matrix subject is overt, the same referent is referred to twice, both in the topic position and in the regular subject position (S *'u*, VX *ka* S). What is the function of this double reference?

The preposed subject is coreferential with the matrix subject. How does the appearance of an NP in preposed position differ from an NP that is the matrix subject, in internal constituency? I will compare the NPs in these cases.

3.3.1 When the matrix subject is overt and when it is not

Among 66 instances of preposed subjects, 13 instances are *S 'u, XV ka hiya* pattern (with overt matrix clause subject), and 53 instances are *S 'u, VX* pattern (without overt matrix subject). The types of preposed NPs are shown in Table 4.

There are considerable differences in internal constituency between *S 'u, VX ka hiya* pattern and *S 'u, VX* pattern.

Table 4: NP types of preposed subjects with overt matrix subjects and those without.

	Preposed subjects			
	S 'u, VX ka hiya.		S 'u, VX.	
PP	0	0%	2	4%
D	1	8%	5	9%
PD	4	31%	8	15%
CD	1	8%	3	6%
BPN	6	46%	14	26%
CG	0	0%	6	11%
CR	0	0%	4	8%
CN	0	0%	4	8%
Saw+ noun	1	8%	6	12%
VP	0	0%	1	2%
Total	13	100%	53	100%

We can first point out that what can appear as S of the *S 'u, VX ka hiya* pattern is very limited. Most of them (10 of 13) are PD or BPN. There is one example each of D, CD and *saw + noun*. There are no examples of PP, CR, CG, CN or VP.

As for the PP (personal pronoun) type, for all preposed subjects there are very few instances, as we saw in §3.2. The two that occur lack an overt subject. Clauses with a personal pronoun as preposed subject and a matrix clause with an overt subject seem to be avoided.

Among the 13 instances of preposed subjects with an overt regular-position subject (*S 'u, VX ka hiya* type), 10 (76%) are instances of PD or BPN; that is, a proper name, with or without a modifying demonstrative, is introduced as the preposed NP and then referred to again in the matrix subject position by a pronoun, as in (48).

(48) *Tiwang niyi 'u, Teruku ka hiya.*
Ciwang PROX PRT Teruku SBJ 3SG
'This Ciwang, she was a Teruku person.'

In the *S 'u, VX* pattern, on the other hand, PD and BPN represent only 41 percent of the cases. This is lower than the number of cases with the *S 'u, VX ka hiya* pattern (76%).

As for the CR, CG, and CN types (common nouns modified by a VP or another noun, those modified by a genitive pronoun, and bare common nouns), there is no instance of the *S 'u, VX ka hiya* pattern. There is only one instance of the CD type (common noun modified by a demonstrative). We can thus say that common nouns rarely appear as preposed subjects of the *S 'u, VX ka hiya* pattern.

For the *S 'u, VX* type, the proportion of common nouns as preposed NPs is higher (17 among 53) than for the *S 'u, VX ka hiya* type.

As for the *saw* + *noun* type, 6 of 7 instances of them occurred in the *S 'u, VX* type.

We see in Table 4 that most instances of *saw* NP *'u* (6 out of 7) occurred in the *S 'u, VX* pattern.

There is one example of VP in preposed position (see the next section for this example).

3.3.2 Preposed subjects and matrix subjects

Now let us compare preposed subjects and matrix subjects. There are 262 overt matrix subject occurrences in the text. A comparison of NP types in preposed subjects and matrix subjects is given in Table 5 and visualized in Figure 2.

Table 5: NP types of preposed subjects with and without overt regular-position subjects and matrix subjects.

	Preposed-Subject				Matrix clause subjects	
	S 'u, VX ka hiya.		S 'u, VX.		XV ka S	
PP	0	0%	2	4%	52	20%
D	1	8%	5	9%	4	2%
PD	4	31%	8	15%	10	4%
CD	1	8%	3	6%	2	1%
BPN	6	46%	14	26%	64	24%
CG	0	0%	6	11%	49	19%
CR	0	0%	4	8%	10	4%
CN	0	0%	4	8%	48	18%
Saw+noun	1	8%	6	12%	2	1%
VP	0	0%	1	2%	21	8%
Total	13	100%	53	100%	262	100%

The PP type (personal pronouns) scarcely appears as a preposed subject, but for matrix subjects, the PP type is observed very often (20%).

Demonstratives (D), proper names modified by demonstratives (PD), and common nouns modified by demonstratives (CD) appear as preposed subjects (22 out of 66, 33%), but not as often in regular subject position (16 out of 262, 6%).

CD, CR, CG (common nouns modified by demonstratives, another noun, a VP, or a genitive pronoun) and CN (bare common nouns) are very common matrix subjects (109 out of 262, 41%). For preposed subjects, they are much fewer (1 out of 14, 7%, for S of the *S 'u, VX ka hiya* pattern and 17 out of 52, 32%, for the S of *S 'u, VX* pattern).

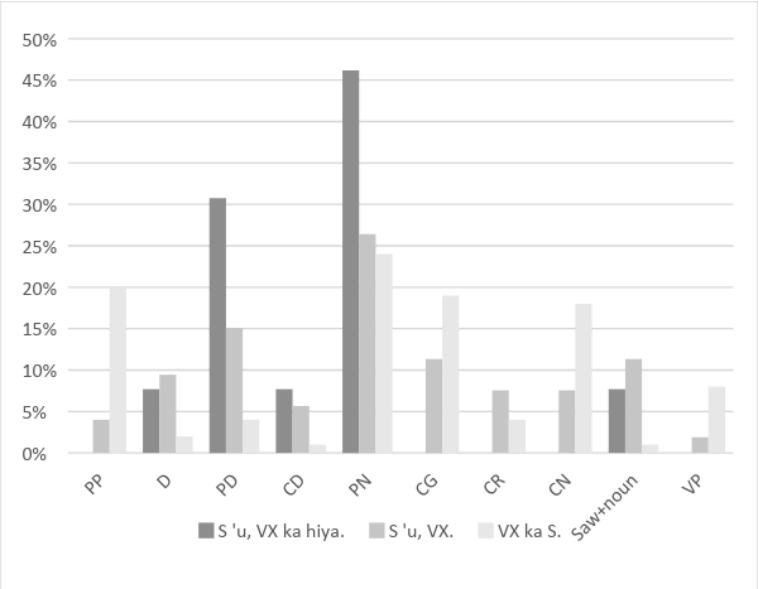

Figure 2: NP types of preposed subjects with and without overt regular-position subjects and matrix subjects.

Personal names, bare (BPN) or with modification by demonstratives (PD), are less common for matrix subjects (74 out of 262, 28%) than for preposed subjects (10 out of 13, 77%, for S of the *S 'u, VX ka hiya* pattern and 22 out of 53, 42%, for S of the *S 'u, VX* pattern).

VPs can function as referential expressions, as we saw at the beginning of §3. Such expressions appear in regular subject position, but rarely in preposed position. There were 21 instances in which a VP appeared as a matrix subject, but only one of a prepositional phrase in pre-clausal position. Example (49) shows a VP functioning as subject.

(49) VP as the subject
Laqi=na *'Ipay Yudaw ka* *q<en>ita.*
child=3SG.GEN Ipay Yudaw SBJ <AV><PRF>see
'The one who saw it was her child Ipay Yudaw.'

We can say that NP *ka* VP is similar to a cleft or pseudocleft construction. Similar constructions are reported for Philippine languages (Himmelmann 2005: 140, Nagaya 2011: 604), and for many other Formosan languages as well.

A preposed VP is usually interpreted as an adverbial-clause predicate with the subject omitted, and not as a referential phrase. This is because *'u* is multifunctional and can mark adverbial clauses as well as preposed subjects, as we saw in §1.4. In the texts, there was one example of a preposed VP that could be interpreted as a referential expression (see example (50)).

(50) VP preceding the matrix clause
P<en>le'alay=bi senehiyi kari Kiristu ka 'alang Besuring 'u, Talug Payan.
<AV.PRF>first=really AV.believe word Christ LNK village Besuring PRT Talug Payab
'As for the first one to believe the gospel of Christ in the village of Besuring, it was Talug Payan.'

The VP preceding *'u* in (50), *p<en>le'alay=bi senehiyi kari Kiristu ka alang Besuring* cannot be interpreted as an adverbial, meaning 'as/because/if one was the first to believe the gospel of Christ in the Besuring village', but only as a referential phrase, meaning 'the first one to believe the gospel of Christ in the Besuring village'.

Instances of *saw* + noun 'such things as *noun*, such things related to *noun*' are very few for matrix subjects (1 out of 262). We can see more examples in preposed position (1 out of 13 for S of the *S 'u, VX ka hiya* pattern, 6 out 53 for S of the *S 'u, VX* pattern). This distribution seems to be the opposite of that with VPs.

4 Anaphoricity

In this section, I will investigate anaphoricity in preposed subjects.

4.1 Preliminaries

Anaphoricity is an index of the degree to which a referent can be said to have a discourse antecedent. Following Gregory & Michaelis (2001: 1687), I apply the label "anaphoricity" to an attribute with three possible values:

- 0: Tokens containing pre-clausal NPs whose referents have not been mentioned in the preceding discourse.
- 1: Tokens containing pre-clausal NPs whose referents are members of a set that was previously evoked.
- 2: Tokens containing pre-clausal NPs that denote entities that have been mentioned previously in the discourse.

Examples (51)–(53) illustrate the three possible anaphoricity scores, with referring expressions and their antecedents co-indexed in (52) and (53).

(51) is an example of an anaphoricity score of 0. The preposed NP in this example is *saw ka qaya samat* 'those things concerning wild animals.' In the previous text, it is not mentioned at all, so its anaphoricity score is 0.

(51) Anaphoricity score of 0 (not mentioned before)
Saw ka qaya samat *'u, des-un=deha be'nux, seberig-an*
like LNK thing wild:animal PRT bring-GV1=3PL.GEN plain:land sell-GV2
kelemukan.
Taiwanese
'As for those things concerning wild animals (animal skins, bones, horns, and the like), they brought them to the plain and sold them to Taiwanese.'

(52) has an anaphoricity score of 1. The preposed NP in example (52b) is *duma* 'some, others.' It is a member of *dehiyaan* '3PL.OBL,' which was mentioned in (52a).

(52) Anaphoricity score of 1 (members of a set that was previously evoked)
a. *Bitaq saw m-ahu quyu peqeraqil dehiya'an*$_{i}$.
until like AV-wash snake AV.torture 3PL.OBL
'They tortured them up to hitting snake (=expression of harshness).'
b. ***Duma***$_{i}$ *'u, pesa-'un=deha kulu m-banah.*
some PRT put-GV1=3PL.GEN box AV-red
'Some were put into prison (=red box).'

(53) has an anaphoricity score of 2. The preposed NP in example (53b) is *dexegal Taiwan* 'land of Taiwan.' It is mentioned in the previous clause (53a).

(53) Anaphoricity score of 2 (Mentioned before)
a. *Saw niyi 'u, berah 'ini 'iyah* ***dexegal Taywan***$_{i}$ *hini ka Nihung*
like this PRT before NEG AV.NFIN.come land Taiwan here SBJ Japan
han.
temporarily
'Things like these, it was before Japan came here to Taiwan.'
b. *Kiya ni pa'ah hengkawas 1895 siida,* ***dexegal Taywan***$_{i}$ *'u diy-un*
so and from year 1895 then land Taiwan PRT have-GV1
kelawa Nihung.
<AV>govern Japan
'Then, from 1895, the land of Taiwan was owned and governed by Japan.'

4.2 Results

For preposed subjects, 35 out of 66 had anaphoricity scores of 2, 1 out of 7 had a score of 1, and 24 of 66 had scores of 0. The results are summarized in Table 6.
In nearly half of the instances the score was 2.

Table 6: Anaphoricity of preposed subjects

	Preposed Subjects	%
0 (=no prior mention)	24	36%
1 (=member of an activated set)	7	11%
2 (=prior mention)	35	53%
Total	66	100%

4.3 Comparison

In this section I will compare the anaphoricity scores of preposed NPs with and without overt matrix subjects and also with matrix subjects.

4.3.1 When the matrix subject is overt and when it is not

A comparison of the anaphoricity of preposed subjects with and without overt matrix subjects is shown in Table 7.

Table 7: Anaphoricity of preposed subjects with and without overt matrix subjects

	Preposed subject			
	S ’u, VX ka hiya.		S ’u, VX.	
0 (=no prior mention)	7	54%	17	32%
1 (=member of an activated set)	0	0%	7	13%
2 (=prior mention)	6	46%	29	55%
Total	13	100%	53	100%

For S of the *S ’u, XV ka hiya* pattern, the anaphoricity score is either 0 or 2; there is no instance of an anaphoricity score of 1. Seven out of 13 are examples where the score is 0, the other 6 are examples where the score is 2. For S of the *S ’u, XV ka hiya* pattern, about a half of the instances (29 of 53) had scores of 2. We can say that the proportion of scores of 2 is similar when the matrix subject is present or absent. For scores of 0 and 1, two cases show a difference: for S of the *S ’u, XV* pattern, the anaphoricity score is 0 (17 out of 53) or 1 (7 out of 53), whereas there are no instances of anaphoricity scores of 1 for S of the *S ’u, XV ka hiya* pattern.

4.3.2 Preposed subjects and matrix subjects

A comparison of the anaphoricity of preposed subjects and matrix subjects is given in Table 8, and illustrated in Figure 3.

Table 8: Anaphoricity of preposed subjects with and without overt matrix subjects, and matrix subjects.

	Preposed subjects				Matrix subjects	
	S 'u, VX ka hiya.		S 'u, VX.		VX ka S.	
0 (=no prior mention)	7	54%	17	32%	98	37%
1 (=member of an activated set)	0	0%	7	13%	15	6%
2 (=prior mention)	6	46%	29	55%	149	57%
Total	13	100%	53	100%	262	100%

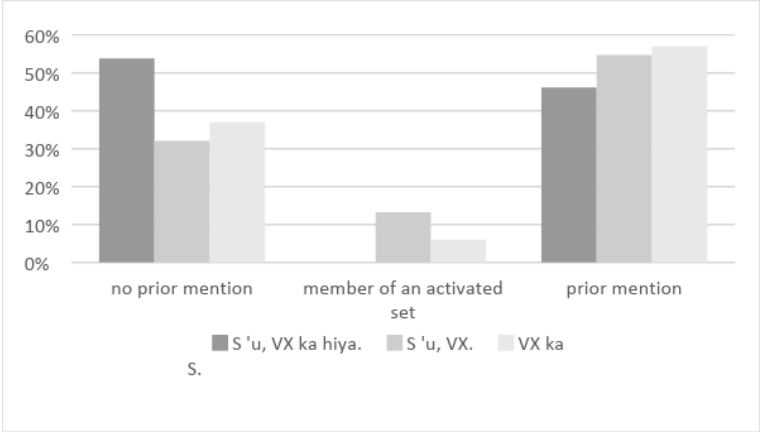

Figure 3: Anaphoricity of preposed subjects with and without overt matrix subjects, and matrix subjects.

For matrix subjects, the proportion of scores of 0, 1, or 2 is somewhat similar to the S of the *S 'u, XV* pattern. For both, more than half of the referents (57% for matrix subjects and 55% for S of the *S 'u, XV* pattern) had anaphoricity scores of 2 (mentioned in the previous discourse), and about one third (37% for matrix subjects and 33% for preposed subjects without overt matrix subjects) had scores of 0 (not mentioned at all). The percentage of NPs with scores of 1 is somewhat lower for matrix subjects (6%) than for S of the *S 'u, XV* pattern (13%).

For S of the *S 'u, XV ka hiya* pattern, the rate of each score differs from the other two types. There is no instance of score 1, and the percentage for score 0 is somewhat higher than for the other two. It may be a coincidence which originates from scarcity of the data.

4.4 Correlation between NP types and anaphoricity

Let us see the correlation between NP types and anaphoricity. It is summarized in Table 9.

Table 9: Type of NP and anaphoricity

	Anaphoricity score		
	0	1	2
PP	0	0	2
D	0	0	6
PD	1	0	11
CD	1	0	3
BPN	12	0	8
CG	4	0	2
CR	2	2	0
CN	1	3	0
Saw+ noun	2	2	3
VP	1	0	0
Total	24	7	35

It is interesting to see the anaphoricity difference between PD and BPN. For most of the PD (11 of 12), the anaphoricity score is 2, while more than half of BPN (12 of 20) had a score of 0. When something already referred to is mentioned again, it appears accompanied by demonstratives. This seems to support the idea that PDs are *activated* or *familiar* while BPN are at least *uniquely identifiable*. For all PP and D also, the anaphoricity score is 2. This seems to support the idea that PP and D correspond to *in focus* and *activated*, respectively.

For those that would correspond to lower givenness (CR, CN, *saw* + noun or VP), we can see that score of 2 is not observed very much.

5 Topic persistence

Lastly, I will examine the topic persistence of preposed subjects.

5.1 Preliminaries

I applied the label of persistence to an attribute with four possible values in the following way:

- 0: The pre-clausal NP denotatum is not referred to at all within five subsequent clauses.

- 1: The pre-clausal NP denotatum is referred to in subsequent clauses by means of a lexically headed NP rather than a pronoun.
- 2: The pre-clausal NP denotatum is expressed pronominally within the five following clauses.
- 3: The pre-clausal NP denotatum is referred to in subsequent clauses by means of a zero pronoun.

I added a fourth value to the classification in Gregory & Michaelis (2001: 1689), because in Seediq one can omit NPs that are recoverable from the context. I wanted to distinguish covert and overt pronouns, so I limited the score of 2 to those cases with overt pronouns and applied a score of 3 to those cases with covert pronouns. The examples in (54-57) illustrate the four possible persistence scores, with referring expressions and their antecedents co-indexed.

The preposed NP of (54a), *Tiwang niyi* 'this Ciwang,' is not referred to at all in at least the following five clauses (though the example below only shows the following two clauses), so the score is 0. From sentence (54b), the topic of the text is changed to *Bakan Hagay* and *Karaw Wacih.*

(54) Lack of persistence; score of 0

a. *Tiwang niyi 'u, dima sedu'uy kari Kiristu ka hiya da.*
Ciwang PROX PRT already AV.have words Christ SBJ 3SG NS
'This Ciwang, she was already a Christian.'

b. *Pa'ah d<en>ehuq-an=na Ekedusan ka Bakan Hagay 'u*
from <PRF>arive-GV=3SG.GEN Ekedusan SBJ Bakan Hagay PRT
'Since Bakan Hagay arrived at Ekedusan,'

c. *'ida sekuxul=bi m-uyas 'uyas 'Utux Baraw deha Karaw Watih senaw=na.*
surely <AV>like=really AV-sing song God Heaven with Karaw Wacih husband=3SG.GEN
'she liked singing songs about God in heaven with her husband Karaw Wacih.'

The preposed NP of (55a), *Lebak Yudaw* 'Lebak Yudaw,' is referred to by a lexically headed NP *Lebak niyi* 'this Lebak' in sentence (55b), the following clause. The score is therefore 1.

(55) Repeated NP; score of 1

a. *Si'ida ka* ***Lebak Yudaw***$_{i}$ *'u, me-'ayung sapah kensat ka hiya ni*
then LNK Lebak Yudaw PRT AV-assistant house police SBJ 3SG and
'At that time, Lebak Yudaw was an assistant at the police station and,'

b. *m-bahang kari quri t<en>egesa Kiristu ka* ***Lebak niyi***$_{i}$ *'uri,*
AV-listen story about <CV>PRF.teach Christ SBJ Lebak PROX also
'this Lebak also heard the story about what Christ taught.'

The preposed NP in (56a), *niyi* 'this,' is referred to by the pronoun *kiya* 'it, so' in sentence (56b), the following sentence. The score is therefore 2.

(56) Pronominal use; score of 2

a. ***Niyi***$_{i}$ *'u, 'adi 'utux=ta ka kiya*$_{i}$*,*
PROX PRT NEG God=1PI.GEN SBJ it
'This, it is not our God,'

b. *'utux 'amirika ka* ***kiya***$_{i}$.
God America SBJ it
'it is the American God.'

The preposed NP of (57a), *kari niyi* 'this story' is the patient of the predicate verb of clause (57b), *m-iyah tegesa* 'come to teach,' but it is not overt. The score is therefore 3.

(57) Zero pronoun use; score of 3

a. ***Kari niyi***$_{i}$ *'u, n-eyah-an=na m-angal pa'ah 'alang Besuring.*
story PROX PRT PRF-come-GV=3SG.GEN AV-take from village Besuring
'As for this story [=the gospel], it was taken from Besuring village.'

b. *Pekelug m-iyah tgesa* __$_{i}$ *hiya ka Tiwang 'Iwal.*
just AV-come <AV>teach there SBJ Ciwang Iwal
'Ciwang Iwal came to teach it just then.'

5.2 Results

For preposed subjects, the topic persistence score is 0 for 35%, 1 for 23%, 2 for 33%, and 3 for 9%. The incidence of score 3 is rather low. It is summarized in Table 10.

Table 10: Persistence of preposed subjects

	Preposed subject	%
0 (=no persistence)	23	35%
1 (=repeated NP)	15	23%
2 (=pronominal use)	22	33%
3 (=zero pronoun)	6	9%
Total	66	100%

5.3 Comparison

In this section I will compare the topic persistence scores of preposed NPs with and without overt matrix subjects, and also with matrix subjects.

5.3.1 When the matrix clause subject is and is not overt

A comparison of topic persistence of preposed subjects depending on the presence or absence of overt matrix subjects is given in Table 11.

Table 11: Persistence of preposed subjects with and without overt matrix subjects

	Preposed subjects			
	S 'u, VX ka hiya.		S 'u, VX.	
0 (=no persistence)	2	15%	21	40%
1 (=repeated NP)	3	23%	12	23%
2 (=pronominal use)	8	62%	14	26%
3 (=zero pronoun)	0	0%	6	11%
Total	13	100%	53	100%

Topic persistence differs considerably depending on the presence or absence of an overt matrix subject.

62% (8 out of 13) of the instances of the S of the *S 'u, VX ka hiya* pattern showed a score of 2, but only 26% (14 out of 53) of the S of the *S 'u, VX* pattern. In contrast, 15% of the instances (2 out of 13) of the S of the *S 'u, VX ka hiya* pattern showed a score of 0, while 40% of the S of the *S 'u, VX* (21 out of 53) showed a score of 0. There was no instance of an overt matrix subject (*S 'u, VX ka hiya*) that showed a score of 3; the six instances of score 3 (11%) are all S of *S 'u, VX*. For S of *S 'u, VX ka hiya*, the proportion of score 2 (62%) is the highest. For S of *S 'u, VX*, the proportion of score 0 (40%) is the highest.

5.3.2 Preposed subject and matrix subject

Now let us compare preposed subjects with matrix subjects. This is shown in Table 12 and Figure 4.

We can see that the matrix subject shows a tendency quite similar to S of *S 'u, VX*. In many instances of both types (45% and 40%, respectively), the referent is referred to only once in a single clause (for a score of 0). In contrast, many instances of S of *S 'u, VX ka hiya* (double reference) show a score of 2 (referred to again by a pronoun within the following five clauses). The doubly referenced entity somehow tends to persist longer.

Table 12: Persistence of preposed subjects with and without overt matrix subjects

	Preposed subjects				Matrix subjects	
	S ’u, VX ka hiya.		S ’u, VX.		VX ka S.	
0 (=no persistence)	2	15%	21	40%	117	45%
1 (=repeated NP)	3	23%	12	23%	51	19%
2 (=pronominal use)	8	62%	14	26%	65	25%
3 (=zero pronoun)	0	0%	6	11%	29	11%
Total	13	100%	53	100%	262	100%

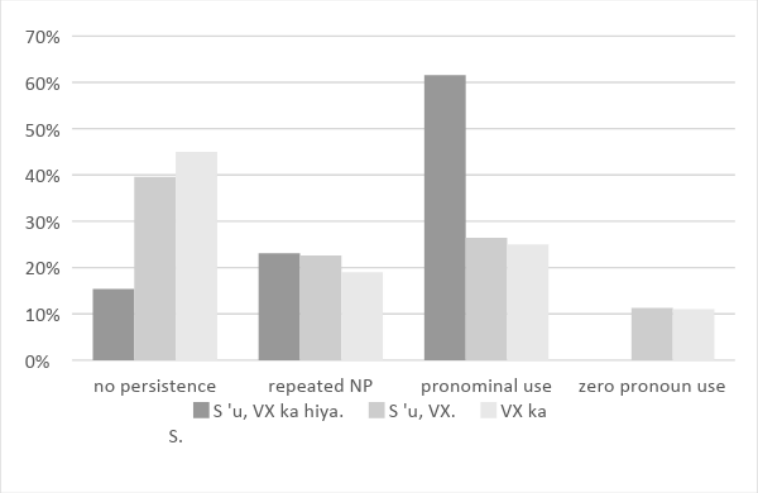

Figure 4: Persistence of preposed subjects with and without overt regular-position subjects

5.4 Correlation between NP types and topic persistence

Let us see the correlation between NP types and topic persistence, summarized in Table 13.

Table 13: NP types and Topic pessistence

	Topic persistence			
	0	1	2	3
PP	0	0	2	0
D	2	1	2	1
PD	2	2	6	2
CD	1	1	1	1
BPN	5	8	6	0
CG	4	0	1	1
CR	2	1	2	0
CN	2	2	0	0
Saw+ noun	4	0	2	1
VP	1	0	0	0
Total	23	15	22	6

As for topic persistence, the tendency is not as clear as in the case of anaphoricity. We can say that CR, CG, CN and *saw* + noun, which are supposed to be at the lower position in the givenness hierarchy of Gundel et al. (1993), the topic persistence score tends to be low (score of 0).

6 Summary

I examined the semantics/function of NP type, anaphoricity, and topic persistence in preposed NPs.

Semantically, more than two thirds of the preposed NPs are coreferential with the matrix subject. One sixth of them are preposed Time.

Preposed subjects, As, and possessors can be regarded to indicate aboutness topics, and the matrix clause denotes the comment about them. Preposed time, alternative, and those preposed NPs that are not classified as any of the above seem to set a frame to the information provided by the matrix clause.

As for NP types, this paper pointed out several characteristics of preposed NPs. We can say that preposed subjects tend to be proper names. About half of the preposed subjects are bare proper names (BPN) or proper names modified by a demonstrative (PD). Only 28% of the matrix subjects are BPN or PD, so BPN or PD is a characteristic of preposed subjects.

Another characteristic of the NP types of preposed subjects is that personal pronouns rarely appear preposed. Only 3% of preposed subjects are personal pronouns, but 20% of matrix subjects are personal pronouns.

We can point out one characteristic of S of the *S 'u, VX ka hiya* pattern: common nouns rarely appear. When the matrix subject is overt, there are very few instances (7%) of common nouns (bare CN), modified by a demonstrative (CD), by another noun or a VP (CR), or by a genitive pronoun (CG). For S of the *S 'u, VX* pattern, one third of preposed subjects are CN, CD, CR, or CG. For matrix subjects, the rate of CN, CD, CR, and CG is higher; it is about 40%.

As for anaphoricity, there is no drastic difference between S of the *S 'u, VX ka hiya* pattern, S of the *S 'u, VX* pattern, or matrix subjects. Matrix subjects showed slightly more (57%) high scores (2) than preposed subjects (53%). Half of the instances of S of the *S 'u, VX ka hiya* pattern (50%) showed a score of 0, which is the highest proportion among the three.

This paper also examined coreference between types of NPs and anaphoricity. Difference in anaphoricity between PD and BPN seems worth noting. Most of the PD are already mentioned in the previous discourse, while BPN are not necessarily so. Also, PP and D are all mentioned in the previous discourse.

As for topic persistence, preposed S of the *S 'u, VX* pattern and matrix subjects showed similar tendencies (Table 12 in §5.3.2). Totally, 40% or 45% showed a score of 0, 23% or 19% showed a score of 1, 25% showed a score of 2, and 12% or 11% showed a score of 3. S of the *S 'u, VX ka hiya* pattern, on the other hand, showed a different tendency: only 14% showed a score of 0, 21% showed a score of 1, and 64% showed a score of 2. There are no instances of preposed subjects with overt matrix subjects that showed a score of 3.

To summarize, we can say that S of the *S 'u, VX* pattern and matrix subjects showed similar tendencies except for NP types. What type of NP it is determines whether it appears in preposed position or in regular subject position. S of the *S 'u, VX ka hiya* pattern, on the other hand, is used to give further information about a proper name. It tends to persist longer in the following clauses. As for anaphoricity, there is no drastic difference among the three.

Acknowledgments

I thank two anonymous reviewers for useful comments. Also I would like to thank Editage (www.editage.jp) for English language editing. For any mistakes in the present paper, however, solely the author is responsible.

Abbreviations

ANGER	anger	CNJ	conjunction
AV	actor voice	CV	conveyance voice
BPN	bare person name	CR	common noun modified by another noun or VP
CG	common noun modified with Genitive pronoun	D	demonstrative
CN	common noun	DIST	distant

EX	exclusive		by demonstrative
FUT	future	PL	plural
GEN	genitive	PO	Possessor
GENTLE	gentle	PP	personal pronoun
GV	goal voice	PRF	perfective
GV1	goal voice 1	PROX	proximant
GV2	goal voice 2	PRT	particle
HEARSAY	hearsay	REASON	reason
IN	inclusive	SBJ	subject
LNK	linker	SG	singular
NEG	negative	UNCERTN	uncertain
NFIN	non-finite	1	first
NOM	nominative	2	second
RDPL	reduplication	3	third
PD	person name modified		

References

Gregory, Michelle L. & Laura A. Michaelis. 2001. Topicalization and left-dislocation: A functional opposition revisited. *Journal of Pragmatics* 33. 1665–1706.

Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274–307.

Himmelmann, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Nikolaus P. Himmelmann & K. Alexander Adelaar (eds.), *The Austronesian languages of Asia and Madagascar*, 110–181. London: Routledge.

Huang, Lillian & Wu Hsinsheng. 2016. *An introduction to atayal grammar*. Shin-pei city, Taiwan: Council of Indigenous peoples.

Krifka, Manfred. 2007. Basic notions of information structure. In C. Fery & M. Krifka (eds.), *Interdisciplinary studies on information structure*, 13–56. Potsdam: Universitätsverlag.

Nagaya, Naonori. 2011. Rise and fall of referentiality: Articles in Philippine languages. In Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages*, 589–626. Philadelphia, PA: John Benjamins Publishing Company.

Tsukida, Naomi. 2009. *Grammar of Seediq, Taiwan*. Tokyo: University of Tokyo (Doctoral dissertation).

Tsukida, Naomi. 2012. Goal Voice and Conveyance Voice of Seediq. In Nakamura & Kikusawa (eds.), *Objectivization and subjectivization: A typology of voice systems* (Senri Ethnological Studies 77), 77–95. Osaka: National Museum of Ethnology.

van Lier, Eva & Jan Rijkhoff. 2013. Flexible word classes in linguistic typology and grammatical theory. In Jan Rijkhoff & Eva van Lier (eds.), *Flexible word classes. Typological studies of underspecified parts of speech*, 1–30. Oxford & New York: Oxford University Press.