TOPICALITY AND REFERENCE-TRACKING IN TAGALOG

Naonori Nagaya
University of Tokyo, Department of Linguistics
nnagaya@gengo.l.u-tokyo.ac.jp

Abstract

The goal of this paper is to demonstrate that, in Tagalog, a pronoun tends to refer to a topical participant, whereas zero anaphora is likely to be employed for an accessible yet non-topical referent, by examining the Tagalog Pear stories. We will also argue that pronominalization serves to express topic continuity in Tagalog: a topic is tracked across clauses by being pronominalized.

1. Introduction

One of the most important functions of language is to track a participant from one clause to the following clauses. This function is called 'reference-tracking' (Foley and Van Valin 1984, Comrie 1989, 1997, 1999). The reference-tracking systems show great diversity: different languages have different reference-tracking devices with different constraints.

This study aims to clarify how Tagalog speakers monitor a specific participant in a discourse. Our main claim is that a topic is monitored by pronominalization: a topic is successively pronominalized from one clause to the succeeding clauses. Tagalog has zero anaphora too, but it is used for an accessible yet non-topical participant. The focus system, which is a set of verbal morphology specific to Tagalog, is irrelevant to this reference-tracking system.

This paper is organized as follows: after we overview Tagalog grammar in Section 2, we will review the typology of reference-tracking systems and some reference-tracking devices in Tagalog in Section 3. In Section 4, we will introduce the Tagalog Pear corpus, which we use in this study. In this section, we will also delimit what we examine in the following section. In Section 5, we will argue, by employing the Tagalog Pear corpus, that a topical participant tends to be marked by a pronoun, and an accessible yet non-topical participant is likely to be referred to by zero anaphora. We will also insist that pronominalization of a topic is an established reference-tracking system in Tagalog. We will conclude this paper with further speculation in Section 6.

This paper is based on a part of my M.A. thesis (Nagaya 2006) submitted to the University of Tokyo. All Tagalog examples employed here are ones I have collected in my field trips to the Philippines.

2. Background

Tagalog is one of the Austronesian languages spoken in the island of Luzon, the Republic of the Philippines. Tagalog is a predicate-initial language, and has an ergative case-marking system (Nolasco 2005).

2.1. Sentence structure

The canonical sentence pattern of Tagalog can be illustrated as below:

(1)   [sentence (L-dislocation) [clause Predicate + Argument(s) + Adjunct(s)]]

The L(eft)-dislocated position is a pragmatically motivated optional position which houses a presupposed element (Nagaya 2005a, to appear).

2.2. Clause members

Like other languages, Tagalog clauses consist of various members (Nagaya 2006).

[1] Predicate. Predicates can be divided into two classes: verbal predicates and nonverbal predicates. Verbal predicates include one of the focus affixes, and nonverbal ones do not. We will discuss the focus system in Section 2.3.

[2] Enclitic. Enclitic members are composed of three subclasses: pronominal enclitics (personal and demonstrative), aspectual enclitics (*na* 'already' and *pa* 'still'), and adverbial enclitics (ex. *lang/lamang* 'only, just', *naman* 'indeed'). Pronominal enclitics inflect for case: absolutive, ergative/genitive, and dative.

[3] Noun phrase and prepositional phrase. Noun phrases are introduced by proclitic case markers. Tagalog has three classes of proclitic markers: absolutive, ergative/genitive, and dative. The absolutive case marks S and O core arguments. (Note that we employ 'S', 'A' and 'O' in Dixon (1972: xxii)'s sense: S = 'subject of an intransitive verb', A = 'subject (or agent) of a transitive verb', and 'O' = 'object of a transitive verb'.) This marker is also employed to encode a left-dislocated element.

The ergative case and the genitive case are formally identical but functionally different. The ergative case marks an A core argument. In contrast, the genitive case is used to encode general possessive relationships and to introduce an adjunct in a clause. Note that the ergative/genitive case marker is spelled as *nang*, not *ng* in this paper. The dative case marks a recipient, a goal, a location and other semantic roles.

Prepositional phrases are employed for peripheral adjuncts.

[4] Complementation and modification. Complementation and modification are both expressed by the linker *-ng/na*.

2.3. Focus system

Nagaya (2006) analyzes the focus system as verbal morphology for equipollent derivation (Haspelmath 1993) or double derivation (Nichols et al. 2004). Both transitive and intransitive verbs are marked by one of the focus affixes: *m-/-um-* (Actor Focus (AF)), *-in* (Patient Focus (PF)), *-an* (Locative Focus (LF)) and *i-* (Circumstantial Focus (CF)). Goal Focus (GF) is used as a cover term for PF, LF and CF verbs.

In general, an AF verb is intransitive, whereas the GF counterpart is transitive. There are two major relationships between intransitive and transitive verbs.

[1] S=O alternation. The S of AF intransitive clauses corresponds to the O of GF transitive clauses.

(2)   a. H<um>into   ang   kotse(S).
          AF:stopped   ABS   car
          'The car stopped.'
       b. I-hinto   mo(A)   ang   kotse(O).
          CF:stop   2SG.ERG   ABS   car
          '(You) stop the car.'

[2] S=A alternation. The S in AF intransitive clauses is equivalent to the A in GF transitive clauses.

(3)   a. D<um>alaw   ako(S)   sa   kaibigan   ko.
          AF:visited   1SG.ABS   DAT   friend   my
          'I visted my friend.'

b. D\<in\>alaw ko(A) ang kaibigan ko(O).
   PF:visited 1SG.ERG ABS friend my
   (ditto)

This type of alternation includes reflexive/middle alternation, where an AF clause means a reflexive/middle event, while the GF counterpart expresses the non-reflexive/non-middle event.

(4)   a. Nag-ahit si George(S) nang bigote.
           AF:shaved ABS George GEN beard
           'George shaved himself.'
        b. \<In\>ahit ni George(A) ang bigote(O).
           PF:shaved ERG George ABS beard
           'George shaved (someone's) beard.'

We do not analyze these alternations as passivization or antipassivization; both AF and GF verbs are of equal morphological complexity (Shibatani 2002).
It is of significance to emphasize that the focus system is a set of verbal morphology particular to Tagalog, and is irrelevant to pragmatic focus (cf. Nagaya 2005a, to appear).

3. Reference-tracking systems

Reference-tracking is to monitor a participant in an on-going discourse. There are a number of devices which are employed to serve this purpose.

3.1. Typology of reference-tracking systems

The first full-scale typology of reference-tracking systems has been proposed by Foley and Van Valin (1984), which list up four types of devices: switch function, switch reference, gender system, and inference.
One of the most popular examples of reference-tracking systems is switch function, where interclausal coreference of two arguments is indicated by voice opposition. For example, in English, a coreferential argument must be in S or A function in control and coordination constructions (Foley and Van Valin 1984: 322).

(5)   a. Fred wants to see Marsha.
        b. *Fred wants Marsha to see [him(O)]
(6)   a. Max persuaded Fred to see Marsha.
        b. *Max persuaded Fred Marsha to see [him(O)]
(7)   a. Oscar went to the store and spoke to Bill.
        b. *Oscar went to the store and Bill spoke to [him(O)]

(5b), (6b) and (7b) are ungrammatical; the coreferential argument is in O function. For the underlying O to be coreferential with the preceding argument, the passive construction must be used as below:

(8)   a. Fred wants to be seen by Marsha.
        b. Max persuaded Fred to be seen by Marsha.
        c. Oscar went to the store and was spoken to by Bill.

Thus, coreference is marked by switching the function of the coreferential argument by passivization. In Dyirbal, antipassivization achieves the same end (Foley and Van Valin 1984: 7.4, Dixon 1994).

3.2. Reference-tracking devices in Tagalog

Although it has not yet been studied from the view-point of reference-tracking, Tagalog has several reference-tracking devices, some of which will be reviewed in this section.

3.2.1. Coreferential markers

In Tagalog, coreference is usually marked within the clause (cf. Hypothesis on local domain in Comrie 1999). The reflexive expression *sarili* 'self' is an example of such coreferential markers.

(9)   P<in>uri   ni   Darling   ang   kanya-ng   sarili.
         PF:praised ERG Darling   ABS   her-LK   self
         'Darling praised herself.'

(10)   Alam   ni   Dodong$_i$ na   [p<in>uri   ni   Joshua$_j$ ang   kanya-ng   sarili$_{*i/j}$].
         know   ERG Dodong   COMP   PF:praised   ERG Joshua   ABS   his-LK   self
         'Dodong$_i$ knows that Joshua$_j$ praised himself$_{*i/j}$.'

Note that the Tagalog reflexive expression must have a clause-mate antecedent. This is the case with the reciprocal expression *isa't isa* 'each other'.

The focus system may be employed to mark coreference within the clause, as mentioned in Section 2.3 [2]. Information about person, animacy and number encoded in nouns and pronouns may contribute to the marking of coreference, although pronouns do not distinguish gender in this language.

3.2.2. Coreferential deletion

In addition to the special markers for coreference, Tagalog also employs coreferential deletion in control and purpose clauses (cf. Hypothesis on extended domain in Comrie 1999). Here are examples of coreferential deletion in control constructions: a coreferential argument in a control clause is omitted or zero-realized.

(11)   H<in>imok   ko   si   Romio   na   [p<um>unta   [S]   sa   school].
         PF:persuaded 1SG.ERG ABS Romio   COMP   AF:go   DAT   school.
         'I persuaded Romio to go to school.'

(12)   H<in>imok   ko   si   Romio   na   [sampal-in   [A]   si   Dodong].
         PF:persuaded 1SG.ERG ABS Romio   COMP   PF:spank   ABS Dodong
         'I persuaded Romio to spank Dodong.'

(13)   H<in>imok   ko   si   Romio   na   [sampal-in   ni   Dodong [O]].
         PF:persuaded 1SG.ERG ABS Romio   COMP   PF:spank   ERG Dodong
         'I persuaded Romio to be spanked by Dodong.'

Also in a purpose clause, an argument can be deleted when it is coreferential with an argument in the main clause.

(14)   T<um>akbo   ang   bata   para   [b<um>ili   [S]   nang   isda].
         AF:ran   ABS   child   for   AF:buy   GEN   fish
         'The child ran so that [the child] would buy fish.'

(15)   K<in>uha   nang   bata   ang   pera   para   [bil-h-in [A]   ang   isda].
         PF:got   ERG   child   ABS   money   for   PF:buy   ABS   fish
         'The child got the money so that [the child] would buy the fish.'

(16)   B<in>ili   nang   bata   ang   isda   para   [kain-in   nang   nanay   niya [O]].
         PF:bought ERG   child   ABS   fish   for   PF:eat   ERG   mother   his/her
         'The child bought the fish so that his/her mother would eat [the fish].'

The coreferential deletion in Tagalog has two important properties. First, there is no restriction on core arguments which can be involved in this operation. Every core argument (i.e. S,

A or O) can be omitted in these constructions, if it is coreferential with a specific argument in the main clause.

Second, the focus system is irrelevant to this process. Compare (12) with (13). The A argument is deleted in (12), while the O is elided in (13). In spite of this difference, the verbs in each control clause both take the same focus suffix *-in*. In this point, the focus system is different from passivization (or antipassivization) of languages with switch function.

As shown above, Tagalog has several reference-tracking devices. What we are going to examine in the following discussions is another reference-tracking system which indicates coreference in more extended domains.

4. Tagalog Pear corpus

In this study, we are going to examine what we call the Tagalog Pear corpus (Nagaya 2006). It is a set of 23 texts of approximately 70 minutes, which I recorded in Metro Manila and Antipolo city. 23 native speakers of Tagalog were asked to retell the story after watching the Pear Film (Chafe 1980).

Let us delimit what we are going to discuss in this paper. The targets of our study are core arguments (i.e. S, A or O) in the corpus. We exclude a clausal complement and a core argument in a relative clause from the targets. A headless relative clause is treated as a lexical noun phrase. Idiomatic expressions like *tapos na* 'that's it' or *okay na?* 'okay?' are also excluded from the sentences or clauses we analyze.

Regarding zero expressions, this study concentrates on what we call 'anaphoric zero' or zero anaphora, which is employed to refer to an entity presupposed in the discourse (Nagaya 2005b, 2006). Tagalog has other zero expressions: zero expressions in control and purpose clauses (Section 3.2.2), deictic zero, impersonal zero (see also Himmelmann 1999), and a gap of left-dislocation. However, we do not count these but focus on anaphoric zero only.

Thus, our Tagalog Pear corpus contains 1415 core arguments. Table 1 represents which grammatical function is referred to by which means of reference.

	Lexical		Pronoun		Zero		Total
	n	%	n	%	n	%	n
A	56	(13.3)	286	(67.9)	79	(18.8)	421
S	215	(35.6)	327	(54.1)	62	(10.3)	604
O	213	(54.6)	66	(16.9)	111	(28.5)	390
Total	484	(34.2)	679	(48.0)	252	(17.8)	1415

Table 1: Reference forms of core arguments in our corpus

5. Reference-tracking by pronominalization

In this section, we aim to clarify how reference-tracking is going on in the Tagalog Pear corpus, from both quantitative and qualitative perspectives.

In Section 5.1, our quantitative study will reveal that a topical referent tends to be encoded by a pronoun, but an accessible yet non-topical one is likely to be referred to by zero anaphora. In Section 5.2, our qualitative study will show that pronominalization functions to express topic continuity in this language. In Section 5.3, it will be argued that this reference-tracking system can mark coreference of two arguments, whatever their grammatical relations might be. In Section 5.4, this reference-tracking system will be evaluated.

5.1. Non-lexical forms: pronoun and zero anaphora

In Tagalog as many other languages, once a participant is introduced into the discourse as a lexical noun phrase, it is in turn referred to by a non-lexical form in the following clauses for the sake of speakers' economy. Namely a participant recoverable from contexts tends to be pronominalized or zero-realized. But pronominalization and zero anaphora behave differently in

this language. Let us describe the differences by examining our corpus.

First of all let us consider Table 2. It reveals that an animate participant tends to be referred to by a pronoun, but an inanimate one by zero anaphora. There are 781 animate participants referred to by non-lexical forms in our corpus, among which 642 (82.2%) examples are pronouns and 139 (17.8%) examples are anaphoric zero. On the other hand, our corpus contains 150 inanimate participants referred to by non-lexical items, of which 37 (24.7 %) examples are pronouns and 113 (75.3 %) examples are anaphoric zero.

	Pronoun		Zero		Total
	n	%	n	%	n
Animate	**642**	**(82.2)**	139	(17.8)	781
Inanimate	37	(24.7)	**113**	**(75.3)**	150
Total	679	(72.9)	252	(27.1)	931

Table 2: Non-lexical forms and animacy

The second point to notice is about grammatical relations. Cross-linguistically, A and S functions tend to be animate and O function is likely to be inanimate (Dixon 1994). Now we have found that animate participants prefer pronouns and inanimate ones zero anaphora. Then, it is not surprising that the S and the A are usually referred to by a pronoun, but the O is likely to be referred to by a zero, among the core arguments realized as non-lexical forms. Let us see Table 3 and Table 4.

	Pronoun		Zero		Total
	n	%	n	%	n
A	**286**	**(78.4)**	79	(21.6)	365
S	**327**	**(84.1)**	62	(15.9)	389
O	66	(37.3)	**111**	**(62.7)**	177
Total	679	(72.9)	252	(27.1)	931

Table 3: Non-lexical forms and grammatical relations I

	Pronoun		Zero		Total
	n	%	n	%	n
Animate A	**286**	**(78.4)**	79	(21.6)	365
Inanimate A	0	(0)	0	(0)	0
Animate S	**301**	**(88.0)**	41	(12.0)	342
Inanimate S	26	(55.3)	21	(44.7)	47
Animate O	**55**	**(74.3)**	19	(25.7)	74
Inanimate O	11	(10.7)	**92**	**(89.3)**	103
Total	679	(72.9)	252	(27.1)	931

Table 4: Non-lexical forms and grammatical relations II

In conclusion, a topical referent (i.e. an animate participant and/or an S or A core argument) tends to be referred to by a pronoun, while a non-topical referent (i.e. an inanimate participant and/or an O core argument) is likely to be marked by zero anaphora.

5.2. Pronoun chain as topic chain

As shown above, a participant high in topicality tends to be referred to by a pronoun in Tagalog. In this section, we will demonstrate that, because of this property, a topic is tracked across clauses by pronominalization in our corpus. Thus, pronoun chain is equivalent to topic chain in this language: following a series of pronouns in the discourse means monitoring a topic. Let us confirm this claim by examining our data in detail.

First of all, let us consider the following paragraph cited from the corpus. Example (17) describes the scene where a young boy found the basket full of pears, took it and ran away. The

young boy is the topic of this paragraph and is encoded by a pronoun across the clauses. Zero anaphora is put into square brackets. Topical pronouns are indicated in **boldface**, so that it is easier to trace the pronoun chain.

(17) (ROM)

a. pagbaba nang bayk, t<in>ign-an **niya(A)** yung prutas(O).
getting.off GEN bike LF:watched 3SG.ERG ABS fruit
'getting off his bike, he watched the fruits.'

b. t<in>ign-an **niya(A)** ang isa(O).
LF:looked.at 3SG.ERG ABS one
'he looked at one (of them)'

c. pagkuha niya nang isa, hindi **siya(S)** na-kontento.
getting his GEN one not 3SG.ABS PF:was.satisfied
'getting one, he was not satisfied.'

d. b<in>uhat **niya(A)** ang isa-ng kahan .. isa-ng kahon(O).
PF:lifted 3SG.ERG ABS one-LK one-LK box
'he lifted up one box.'

e. d<in>ala **niya(A)** [O] ngayon sa bayk.
PF:carried 3SG.ERG now DAT bike
'he carried [the box] to his bike.'

f. s<in>akay **niya(A)** [O].
PF:put 3SG.ERG
'he put [the box].'

g. di .. <um>alis na **siya(S)** ngayon.
filler AF:left already 3SG.ABS now
'mmm... he left now.'

This situation can be represented as the referent transition diagram in Figure 1.

	boy	fruits	box
a	pro(A)	lex(O)	
b	pro(A)	lex(O)	
c	pro(S)		
d	pro(A)		lex(O)
e	pro(A)		zero(O)
f	pro(A)		zero(O)
g	pro(S)		

Figure 1: The referent transition diagram for (17)

The referent transition diagram clearly shows that the pronouns keep referring to the boy, i.e. the topic of this paragraph. The topic chain, namely, coreferential relationship between the topical arguments is expressed by successive pronominalization of the topic. Zero anaphora, on the other hand, refers to the accessible yet non-topical participant.

Of course, the choice of a topic is flexible: different participants may be selected as a topic in different segments of the discourse. Let us consider (18), where the speaker describes what happened to the boy and the three children in the anterior half, and she is trying to explain the psychology of the boy in the posterior half. Both parts describe substantially the same situation, but the speaker selects the child(ren) as a topic in the former part, and the boy in the latter part. In either part, the topic is pronominalized.

(18) (MRN)

a. paglayo nung isa-ng bata-ng [nasa bayk],
going.away GEN one-LK child-LK was.at bike
na-kita nang tatlo-ng lalaki(A) ang sombrero(O).
PF:found ERG three-LK man ABS hat

          'When the one child riding a bike went away, the three men found the hat.'

b.   <in>akala   **nila(A)**-ng,
       PF:thought   3PL.ERG-COMP
       'they thought ...'

c.   sa   bata, sa bata-ng   lalaki-ng iyon-ng   [naka-bayk]   [S].
       DAT child DAT young-LK   man-LK   that.ABS-LK   AF:rode.bike
       '[the hat] belongs to the child, the boy riding a bike.'

d.   at   muli **niya(A)**-ng t<in>awag [O].
       and   again 3SG.ERG-LK PF:called
       'he called [the boy] again.'

e.   s<in>ipul-an **niya(A)**   [O].
       PF:whistled   3SG.ERG
       'he whistled [the boy].'

f.   at   na-pa-lingon   ang bata-ng   lalaki na [nasa   bayk](S).
       and PF:looked.back ABS young-LK man   LK was.at bike
       'and the boy riding a bike looked back.'

g.   i-s<in>auli **niya(A)** ang sombrero(O).
       PF:returned 3SG.ERG ABS hat
       'he returned the hat.'

h.   at   hindi nag- .. hindi nag-dalawang-isip ang bata-ng   lalaki   na
       and   not   not   AF:hesitated   ABS young-LK man   LK
       [nasa   bayk](S) na big–
       was.at bike
       'and the boy riding a bike did not hesitate.'

i.   bilang kapalit   nang   pagtulong sa   kanya   at pagsauli   sa
       as   exchange GEN   help   DAT 3SG.DAT and returning   DAT
       kanya-ng sombrero, b<in>igy-an siya   nang ..
       his-LK   hat   LF:gave   3SG.ABS   GEN
       b<in>igy-an [A] **sila(O)** nang tagiisa-ng prutas, ay bunga.
       LF:gave   3PL.ABS GEN   each-LK   fruit   fruitage
       'as exchange for the help to him and returning back of his hat, [the boy] gave him– [the
       boy] gave them fruits, no, fruitages.'

j.   hindi **ko(A)**   alam,
       not   1SG.ERG know
       'I don't know...'

k.   kung   prutas o   bunga [S].
       whether fruit   or   fruitage
       'whether [the pear] is a fruit or fruitage.'

l.   b<in>igy-an [A] **siya(O)**.
       LF:gave   3SG.ABS
       '[the boy] gave him.'

m.   ngayon a= sabay   l<um>ayo   na   ang tatlo-ng lalaki(S).
       now   at.the.same.time AF:went.away already ABS   three-LK man
       'now, the three men went away.'

n.   <um>alis na   sabay   din ang bata-ng   lalaki na,
       AF:left   already at.the.same.time too ABS   young-LK man
       bata-ng   lalaki na [nasa   bayk](S).
       young-LK man   LK was.at bike
       'the boy, the boy riding a bike also left together at the same time.'

o.   habang **siya(S)** 'y   papalayo,
       while   3SG.ABS INV going.away
       'while he was going away,'

p.   parang nag-i-isip   **siya(S)** na,
       seem   AF:was.thiking 3SG.ABS COMP
       'it seems that he was thinking ..'

q. sa   kabila   nang   [g\<in\>awa   niya],   k\<in\>uha   **niya(A)**   lang
   DAT   side   GEN   PF:did   3SG.ERG   PF:got   3SG.ERG   just
   yung   prutas(O)   na   ay   bunga   na   [hindi   naman   sa   kanya].
   ABS   fruit   LK   ay   fruitage   LK   not   indeed   DAT   3SG.DAT
   'on the one side of what he did, he just took the fruits or fruitages which did not belong
   to him.'

r. tapos   sa   kabila   pala   noon,   mayroon   pala-ng   ano   manga
   then   DAT side   surprise   that.GEN   be   surprise-LK   filler PL
   bad .. manga–
   bad PL
   'then on the other side, there was something bad.'

s. parang   na-konsiyensiya   **siya(S)**   sa   [g\<in\>awa   niya]   na–
   seem   PF:felt.guilty   3SG.ABS   DAT PF:did   3SG.ERG
   'it seems that he felt guilty at what he did.'

t. hindi   **niya(A)**   lang   na-sabi   [O].
   not   3SG.ERG   just   PF:could.not.say
   'he could not say [it].'

u. kaya   bini-   b\<in\>igy-an   **niya(A)**   [O]   nang   tatlo-ng–
   so   LF:gave   3SG.ERG   GEN   three-LK
   'so he gave [them] three...'

v. bilang   kapalit,   b\<in\>igy-an   **niya(A)**   nang   tatlo-ng   prutas   ang ..
   as   compensation   LF:gave   3SG.ERG   GEN   three-LK   fruit   ABS
   yung   manga   bata(O).
   ABS   PL   child
   'by way of compensation, he gave the chilren three fruits.'

The referent transition diagram for (18) is below.

	speaker	children	child	boy	hat	other inanimates
a		lex(A)			lex(O)	
b		pro(A)				
c					zero(S)	
d			pro(A)	zero(O)		
e			pro(A)	zero(O)		
f				lex(S)		
g			pro(A)		lex(O)	
h				lex(S)		
=>i		pro(O)		zero(A)		
j	pro(A)					
k					zero(S)	
=>l		pro(O)		zero(A)		
m		lex(S)				
n				lex(S)		
o				pro(S)		
p				pro(S)		
q				pro(A)		lex(O)
r						lex(E)
s				pro(S)		
t				pro(A)		zero(O)
=>u		zero(O)		pro(A)		
=>v		lex(O)		pro(A)		

Figure 2: The referent transition diagram for (18)

Compare (18i)/(18l) with (18u)/(18v). All these clauses mean that the boy gave pears to
one of the three children. However, (one of) the three children is(are) a topic and pronominalized

in (18i) and (18l), while the boy is a topic and pronominalized in (18u) and (18v). Note also that zero anaphora serves to disambiguate the topic, by encoding the non-topical referents.

As such, our qualitative studies based on the Pear stories have revealed that pronominalization of a topic, coupled with zero anaphora, serves to track a topic.

5.3. Non-prototypical coreferential patterns

In the preceding examples, the pronominalization is employed to mark the coreferential relationships of S/A patterns or neutral patterns (Tsunoda 1986, in preparation). That is, the pronoun chains link two core arguments of the identical grammatical relation, or A and S core arguments (see Figure 3).

It is not surprising that topic chain prefers S/A patterns and neutral patterns, because "[...] in every language, discourse is organised about a series of 'topics', which are most often human, and controllers of actions, and thus most likely to be in underlying A or S functions [...]" (Dixon 1994: 174).

However, pronoun chain in Tagalog can indicate coreference of S/O patterns and aberrant patterns, too, although it is rare. To put it differently, this reference-tracking system can mark coreference of two arguments, whatever their grammatical relations might be.

	preceding clause	following clause
neutral patterns	S	S
	A	A
	O	O
S/A patterns	S	A
	A	S
S/O patterns	S	O
	O	S
aberrant patterns	A	O
	O	A

Figure 3: Four patterns of coreference

Let us consider example (19) below. In this portion of discourse, the topic is the boy who stole the basket and he is monitored across clauses by pronominalization. What we should pay attention to is the last three clauses.

(19) (MRC)
a. sa ano, yun nga, na-karma siguro **siya(S)**
DAT what that.ABS indeed PF:got.bad.karma probably 3SG.ABS
dahil sa [g<in>awa niya]-ng pagnanakaw.
because.of DAT PF:did 3SG.ERG-LK stealing
'...that's it, probably, he got bad karma because of what he did, stealing.'
b. na-tumba **siya(S)** doon na ano yung bayk niya na na–
PF:fell.down 3SG.ABS there ABS bike his
'he fell down there ... his bike.'
c. na-disgrasya **siya(S)**.
PF:had.misfortune 3SG.ABS
'he had a misfortune.'
d. tapos may manga, may manga bata doon na-
then be PL be PL child there LK
tatlo-ng bata-ng lalaki na [t<um>ulong sa kanya],
three-LK young-LK man LK AF:helped DAT 3SG.DAT
'then there were children there ... three young men who helped him,'

e. pagkatapos   niya-ng   ma-dapa,
   after   his-LK   PF:stumble
   'after he stumbled,

f. o ma-semplang   yung   bayk   niya(S).
   or PF:fall.down ABS   bike   his
   'or his bike fell down,

g. tapos noong   ano, t<in>ulung-an na   [A]   **siya(O)**.
   then   when   what LF:helped   already   3SG.ABS
   'then, when .., [the three children] helped him.'

h. e=   di   tapos   na.
   filler filler finished   already
   'it's finished.'

i. <um>alis   na   yung   tatlo-ng   bata(S).
   AF:left   already ABS   three-LK   child
   'the three children left.'

j. e= .. yung   bata, na-kuha   yung   sombrero niya(O).
         ABS   child PF:got   ABS   hat   his
   '... the child, (he) got his hat.'

k. kasi   may .. yun na,   na-iwan   **niya(A)**   [O].
   because   be   PF:left   3SG.ERG
   'because ... he left [the hat].'

l. tapos doon, siguro   yung   bata na, ano   siya,
   then   there probably ABS   child   what   3SG.ABS,
   na-tuwa   **siya(S)**   doon.
   PF:got.happy   3SG.ABS   there
   'then there probably the child, he got happy at it.'

m. dahil   nga,   dahil   nga   t<in>ulung-an   [A]   **siya(O)**.
   because indeed   because indeed LF:helped   3SG.ABS
   'because, because [the children] helped him.'

n. b<in>igy-an **niya(A)**   nang   tatlo-ng   prutas isa-isa,
   LF:gave   3SG.ERG   GEN   three-LK   fruit   one-by-one
   .. yung   isa-ng   bata-ng   [l<um>apit].
      ABS   one-LK   child-LK AF:came.close
   'he gave three fruits one by one .. to the child who came close.'

Here is the referent transition diagram for (19).

	boy	children	child	bike	hat
a	pro(S)				
b	pro(S)				
c	pro(S)				
d		lex(E)			
e					
f				lex(S)	
g	pro(O)	zero(A)			
h					
i		lex(S)			
j			lex(L)		lex(O)
k	pro(A)				zero(O)
=>l	pro(S)				
=>m	pro(O)	zero(A)			
=>n	pro(A)		lex(O)		

Figure 4: The referent transition diagram for (19)

The topic of this passage is the boy who was helped by the three children. The

grammatical relation of the topic being tracked changes from S to O in (19m), which is an S/O pattern. Then, it turns from O to A in (19n), which is an aberrant pattern. In spite of the differences in grammatical relations, coreference of the topic is marked by pronominalization across the three clauses.

It is noteworthy to point out that this is not the case with switch function. For example, let us look at (19m); the O pronoun refers back to the S pronoun in the preceding clause. This S/O pattern of coreference is disfavored in languages with accusative syntax (Tsunoda 1986, in preparation), and would be avoided by the passive to maintain an S/A pivot.

Here is another example.

(20) (MRT)
    a.   tapos,   noong   nag-la-lakad   **sila(S)**,
        then   while   AF:were.walking   3PL.ABS
        'then while they were walking,'
    b.   na-pa-daan   **sila(S)**   doon   sa   .. isa-ng .. yung   lalaki-ng
        PF:passed.by 3PL.ABS there DAT   one-LK ABS   man-LK
        [na-mi-mitas   nang   peras mula sa   puno].
        AF:was.picking   GEN   pear   from DAT tree
        'they accidentally passed by the man who was picking fruits from the tree.'
    c.   am= tapos, tapos, na-kita   **sila(O)**   nang   lalaki-ng ito   na
            then   then   PF:saw 3PL.ABS   ERG   man-LK   this.ABS LK
        [p<um>i-pitas   sa   puno](A).
        AF:was.picking   DAT tree
        'then this man picking at the tree saw them.'
    d.   at .. iyon,   p<in>ag-masd-an lang   [A]   **sila(O)**.
        and   that.ABS LF:stared   just   3PL.ABS
        'well, [the man] stared at them.'
    e.   tapos, yung   manga bata, hindi   **nila(A)**   na-kita
        then   ABS   PL   child   not   3PL.ERG PF:saw
        yung   lalaki-ng [p<um>ipitas   nang   peras](O).
        ABS   man-LK   AF:was.picking   GEN   pear
        'then, as for the children, they did not see the man picking pears.'
    f.   d<um>aan   lang   talaga **sila(S)**   doon.
        AF:passed.by   just   really   3PL.ABS   there
        'they just really passed by there.'

The referent transition diagram for (20) is given in Figure 5. Notice again that coreference is coded by the pronoun chain, in whatever function the pronouns might be.

	children	farmer
a	pro(S)	
b	pro(S)	
=>c	pro(O)	lex(A)
d	pro(O)	zero(A)
=>e	pro(A)	lex(O)
f	pro(S)	

Figure 5: The referent transition diagram for (20)

5.4. Characteristics of this system

As demonstrated above, pronominalization is one of the reference-tracking devices in Tagalog, and serves to track a participant across clauses. This reference-tracking system has three important characteristics. First, a topic can be tracked across clauses regardless of its grammatical

relations. In our corpus, a topic is prototypically in S or A function, but may be in O function. In any case, this system can mark coreference, as demonstrated in Section 5.3. This contrasts with English reference-tracking by switch function, where an argument being tracked is restricted to S or A function (see again Section 3.1).

Second, the focus system is not relevant to this reference-tracking system. What is relevant is not the verbal form but the reference form of a participant. This implies that the focus system in Tagalog is different from active-passive voice opposition or ergative-antipassive voice opposition.

It is of significance to point out that these two characteristics of the reference-tracking system are parallel to those of coreferential deletion in control and purpose clauses in this language. Coreferential deletion in these constructions has no restriction on core arguments which can be involved in the operation, and the focus system is irrelevant to it (see Section 3.2.2).

Third, pronominalization rather than zero anaphora marks a topic. In the literature, it has been considered that zero anaphora is the least marked coding for a topical referent (Givón 1983). In many languages like Japanese, Korean, and Mandarin Chinese, topic chain is formed by zero anaphora. Although pronominalization is also available in these languages, it is more marked than zero anaphora for a topic. But in Tagalog, a pronoun is employed to mark a topic, while zero anaphora is used for a non-topical participant.

6. Conclusion

In this paper, I have tried to argue that, in Tagalog, a topical participant is likely to be encoded by a pronoun, while an accessible yet non-topical one tends to be referred to by zero anaphora, and pronominalization of a topic, coupled with zero anaphora, serves to express topic continuity in this language. This reference-tracking system has three important characteristics. First, there is no restriction on core arguments which can be involved in this system. Second, the focus system does not contribute to the marking of coreference. Third, a pronoun rather than zero anaphora is employed to mark a topic in this reference-tracking system.

There is one question that has yet to be answered: how come a topical participant is referred to by the more marked expression (i.e. pronoun), and a non-topical one is expressed by the less marked means (i.e. zero)? This might be a compensatory system. English, on the one hand, has a mechanism for switch function, that is, passivization. In this language, zero anaphora is the least marked coding for a topical referent. Tagalog, on the other hand, lacks such a mechanism, but employs a pronoun instead to mark a topic.

Another relevant fact is that the pronouns examined here are basically pronominal enclitics: this reference-tracking system may indicate that Tagalog pronominal system is at the intermediate stage between optional pronominal clitics and obligatory pronominal agreement (Givón 1976, 2001: Chapter 9, Himmelmann 1999).

Abbreviations

The following abbreviations are employed in this paper: ABS-absolutive, AF-actor focus, CF-circumstantial focus, DAT-dative, ERG-ergative, EXC-exclusive, GEN-genitive, INC-inclusive, INV-inversion, LF-locative focus, LK-linker, PF-patient focus, PL-plural, SG-singular, 1-first person, 2-second person, 3-third person, and < >-infix.

Here is the list of abbreviations for the referent transition diagram: boy-the boy who steals the pears, child-the child who picks up the hat of the boy, children-the children who help the boy, farmer-the man who harvests pears, lex-lexical noun phrase, pro-pronoun, zero-anaphoric zero, A/S/O-grammatical relations, E-participant introduced by an existential construction, and L-left-dislocated participant.

Acknowledgements

References

Chafe, Wallace L. 1980. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production.* Norwood, NJ: Ablex.

Comrie, Bernard. 1989. Some general properties of reference-tracking systems. In D. Arnold, M. Atkinson, J. Durand, C. Grover, and L. Sadler, eds., *Essays on grammatical theory and universal grammar*, 37-51. Oxford: Clarendon Press.

1997. Pragmatic binding: Demonstratives as anaphors in Dutch. *Proceedings of the Annual Meeting of the Berkeley Linguistic Society* 23: 49-61.

1999. Reference-tracking: Description and explanation. *Sprachtypologie und Universalienforschung (STUF)* 52: 335-346.

Dixon, R.M.W. 1972. *The Dyirbal language of North Queensland.* Cambridge: Cambridge University Press.

1994. *Ergativity.* Cambridge: Cambridge University Press.

Foley, William A., and Robert D. Van Valin Jr. 1984. *Functional syntax and universal grammar.* Cambridge: Cambridge University Press.

Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. In C. Li, ed., *Subject and topic*, 149-88. New York: Academic Press.

1983. Topic continuity in discourse: An introduction. In T. Givón, ed., *Topic continuity in discourse: A quantitative cross-language study*, 1-42. Amsterdam and Philadelphia: John Benjamins.

2001. *Syntax: An introduction,* vol. 1. Amsterdam and Philadelphia: John Benjamins.

Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In B. Comrie and M. Polinsky, eds., *Causatives and transitivity*, 87-120. Amsterdam: John Benjamins.

Himmelmann, Nikolaus P. 1999. The lack of zero anaphora and incipient person marking in Tagalog. *Oceanic Linguistics* 38:231-269.

Nagaya, Naonori. 2005a. Information structure and constituent order in Tagalog. Paper presented at the 2005 International Course and Conference on Role and Reference Grammar, held at Academia Sinica, Taipei.

2005b. Zero anaphora in Tagalog. Presentation at the Department of Linguistics, University of

the Philippines Diliman on October 1. Quezon city: University of the Philippines Diliman.

2006. *Grammatical relations and reference-tracking systems in Tagalog*. M.A. thesis. University of Tokyo.

To appear. Information structure and constituent order in Tagalog. *Language and Linguistics*. Taipei: Institute of Linguistics, Academia Sinica.

Nichols, Johanna, David A. Peterson, and Jonathan Barnes. 2004. Transitivizing and detransitivizing languages. *Linguistic Typology* 8:149-211.

Nolasco, Ricardo Ma. 2005. What ergativity in Philippine languages really means. Paper presented at Taiwan-Japan Joint Workshop on Austronesian Languages, held at National Taiwan University, Taipei.

Shibatani, Masayoshi. 2002. Gengo-ruikeiron-to taishou-kenkyuu [Linguistic typology and contrastive study]. In N. Ogoshi, ed., *Taishou gengogaku [Contrastive linguistics]*, 11-48. Tokyo: University of Tokyo Press.

Tsunoda, Tasaku. 1986. Topicality in ergative and accusative languages. *Nagoya Working Papers in Linguistics* 2:173-258. Nagoya: Department of Linguistics, Faculty of Letters, University of Nagoya.

In preparation. *The Warrungu language of North Queensland* (tentative title).