# Split O in Formosan Languages—A Localist Interpretation[*]

## Shuanfan Huang
### *National Taiwan University*

The non-agent focus constructions in Formosan languages can be fruitfully interpreted as a species of the applicative construction which functions to derive transitive clauses from intransitive or transitive clauses. The O argument NPs in these transitive clauses are shown to exhibit a split O phenomenon based principally on the nature of verbal semantics and secondarily on discourse-pragmatic considerations: the nominative NP of a PF clause encodes a patient object, that of the LF clauses an abstract location, and that of the RF clauses a transported theme. Other functions of RF clauses are shown to be secondary to and derivable from this more spatial notion via pragmatic inferencing. This paper treads new ground by integrating the semantics of focus into cognitive grammar and typological approaches. Thus all of the coding patterns for the O arguments together form a semantic space, which then constrains possible coding patterns for language-specific constructions, and allows for predictions about the interaction of these applicative clauses.

Key words: split O, focus, Formosan languages, semantic space

## 1. The problem

A problem in the syntax of Formosan languages that has captured the interest of Formosan grammarians for the last two decades has been the nature of what has misleadingly gone under the general name of 'focus constructions'. These 'focus constructions' contain affixes that function to derive transitive clauses from intransitive or transitive clauses and can thus be viewed as applicative constructions (cf. Ross 2002, Croft 2003). In these applicative clauses, the agentive NP is an A and the nominative NP an O, to use the terminology first suggested by Dixon (1972) and Comrie (1978). The O arguments of transitive verbs are expressed in three grammatically distinct ways, depending on the semantics of the verbs to which the applicative affix is attached. Some

---

of the O arguments must be expressed as the nominative NPs of the Patient focus construction (PF); others as nominative NPs of the locative focus construction (LF), still others as the nominative NPs of the Referential focus construction (RF). I suggest that viewing the 'focus constructions' as argument-enhancing strategies makes it possible for us to ask the right sort of question about transitivity. An important question that naturally arises is: What motivates the distinction of the three types of O arguments? A number of researchers have sought to explain the bases for the distinction in usage between the various focus constructions (e.g. Ferrell 1979, Szakos 1994, L. Huang 1995, Yeh 2003). In this paper I tread new ground by proposing that these argument-enhancing constructions exhibit what I shall term the split O phenomenon. I shall demonstrate that the semantic basis for the split O phenomenon can be given a fairly straightforward localist interpretation, though the basis apparently varies from language to language, because of the demonstrable conceptual fluidity and interpenetrability of the various types of O arguments.

## 2. Complexity of the functional differences underlying the PF/LF distinction

An investigation of the distribution of verbs in the various applicative constructions with respect to their co-occurring nominative NPs must begin with the important observation by Starosta (2002) that it is very rare to find a verb in Western Austronesian languages that has forms for all the focus slots. Moreover, the distribution is not amenable to any of the usual functional explanations. It is largely determined by the semantics of the verb, the nature of O arguments, and the presence or absence of a causative element in the verb and has nothing to do with topicality of the subject. Moreover, a proper understanding of the distribution of these different types of verbs can be achieved only by investigating those languages where PF, LF, and RF clause types are still strongly represented. Kavalan, in this view, must be taken out of consideration, since verbs in PF form in this language have merged into LF verbs (a case of analogical leveling—similarity in meaning, or closeness in conceptual space is a cause for similarity in form), and its RF clauses have all but disappeared. Merger between PF and LF forms seems to be widely observable in Formosan languages. In Thao, the affixes for PF and LF forms 'have begun to lose any distinguishing syntactic or semantic characteristics and have become largely interchangeable.' (Blust 2003:92). A check through Egerod's (1999) dictionary, for example, shows that 9% of the verbs in Squliq Atayal appear only in PF forms and 41% of the verbs appear only in LF forms and the remaining 50% appear in both PF and LF forms. Note that Egerod did not indicate which verbs can or must occur only in RF form. Although many Formosanists have generally taken it as a

truism that the distinction between the PF/LF form of a verb corresponds to a difference in aspect, verbs in PF form being used to indicate irrealis, while those in LF form realis, the relationship between the PF/LF forms and their functions is considerably much more complex and murky than commonly believed. There are verbs that occur only in PF or in LF or even only in RF form and so the realis/irrealis distinction cannot apply to them at all. There are also verbs appearing in both PF/LF forms, but are used for a wide range of functions, only one of which being to indicate mood distinctions (realis vs. irrealis). A careful analysis of the functional distinctions marked by PF or LF affixes at least yields the following categories (worked out in collaboration with Maya Yeh):

(1) Functional differences between LF and PF verbs in Squliq
    a. Denoting the realis (LF) vs. irrealis (PF). This is the only function generally acknowledged by most Formosan researchers. E.g. niq-an/niq-un 'to eat' (cf. (1c))
    b. Both forms denoting the irrealis aspect. E.g. stmaq-un/stmaq-an 'to mash; to rot'
    c. Denoting the opposite of the distinction made in (1a). E.g. lpg-un (realis) /lpg-an (irrealis) 'to count'
    d. Denoting the animacy (PF)/inanimacy (LF) distinction. E.g. gb-un 'to embrace sb.'/gb-an 'to embrace st.'; cbiN-un 'to hold an animate object' vs. cbiN-an 'to hold an inanimate object'
    e. Denoting a perfective event (PF) vs. an imperfective event (LF). E.g. nbw-un 'to finish drinking' vs. nbw-an 'to drink at'
    f. Verbs appearing only in PF form (9% according to Egerod 1999). E.g. baq-un 'to know'; p-qas-un 'to feel happy for sb.'
    g. Verbs appearing only in LF form (41% according to Egerod 1999). E.g. pruŋ-an 'to break; to become broken'; sbhoy-an '(wind) to blow; to boil'; s-p-qasan 'to make sb. happy'; p-nbw-an 'let sb. drink'.

The functional distinctions listed above, together with the interpenetrability of the various types of O arguments, to be discussed in detail below, suggest that there is a strong discourse basis for the complex interplay between these two focus forms, auxiliary elements and functions, and the issue must be addressed in each language separately (cf. Huang & Huang 2003 for a detailed investigation of the auxiliary system and the status of reality marking in Tsou).

Distribution patterns of the various clause types vary from language to language. Analyses of the distribution of AF, PF, LF, and RF main clauses in the NTU spoken corpus of a number of Formosan languages and Tagalog yield the following percentage figures.

**Table 1: Distribution of various main clause types in Formosan languages and Tagalog**
(based on NTU spoken corpus of Formosan languages, unless otherwise indicated)

|            | AF         | PF         | LF        | RF/BF/IF  |
|------------|------------|------------|-----------|-----------|
| Tsou*      | 57.5       | 27.2       | 12.3      | 2.9       |
| Squliq**   | 48.2(699)  | 29.9(434)  | 17.7(257) | 4.2(61)   |
| Saisiyat***| 77         | 19.2       | 0         | 3.8       |
| Seediq     | 66.3       | 15.8       | 15.6      | 2.5       |
| Tagalog    | 56.5       | 31.7       | 10.3      | 1.4       |

*In Tsou, verbs with the BF suffix -*neni* very rarely take an instrumental nominative NP.
**Based on the Formosan Language Archive (Academia Sinica), a corpus on Formosan languages.
***In Saisiyat, LF verbs always require an aspect marker *in-* or *ka-* and occur only in subordinate clauses (relative or complement clauses or nominalizaions denoting the Location where an event or activity takes place).

The numbers given in Table 1 show that clauses whose nominative NPs bear the semantic role of location (LF) are far more common than those bearing the role of benefactive or instrument (RF/BF/IF). For some unknown socio-historical reasons, LF verbs in Saisiyat are now, based on our corpus data, reduced to occurring only in subordinate clauses such that verbs that in other languages must occur in LF form when they take an O argument (e.g. perception verbs) now take the PF-marking suffix -*en* (e.g. kita-en 'to see'). Tagalog appears to be pursuing a similar path of development, since it has the lowest percentage of LF verbs and the perception verbs in the language also take the PF affix *na-* (e.g. nakita 'to see-PF'). One notable difference between Saisiyat and Tagalog in this regard is that the loss or decline of verbal LF clauses does not seem to have resulted in a net gain for PF clauses in Saisiyat, whereas it has in Tagalog, as can be determined by looking at the percentage figures for PF clauses in these two languages.

The distinction between verbs in PF and those in LF forms in Formosan languages appears to correlate significantly with the Fillmorean distinction between change-of-state (e.g. verbs like *break*) vs. surface contact verbs (e.g. *hit*), though verbs of the latter type must be understood in some suitably extended, abstract sense for us to appreciate the nature of LF verbs in these languages. Thus the LF verbs in most Formosan languages usually take their O arguments that are construable as either (l) Location; (2) Patient; (3) Percept (of a perceptual event); (4) Goal (of a motion or contact event); or (5) Experiencer (of an emotion). The Patient is easily construable as a kind of location (cf. *to kiss someone* vs. *to give someone a kiss*); the Percept is the location toward which a perceptual event is directed; and the Experiencer is the locus of an emotion. What unites these categories is apparently the umbrella abstract notion LOCATION. These suggest that LF verbs in Formosan languages appear to lexicalize the delivery of some force or

entity, either physical or non-physical, to a particular abstract LOCATION. This explains their tendency to allow this abstract LOCATION to be encoded as the nominative NP of the LF clauses. By contrast, the nominative NPs of the PF clauses are construed as referring to OBJECTs. An important corollary of this way of conceptualizing the events associated with LF clauses is the interpretation of states as metaphorical LOCATIONs perhaps through recurring instances of a particular emotional state being experienced in a specific locale, an interpretation that comports with the discussions over the last twenty years in cognitive linguistics. **Locative expressions also normally cover the domain of time as well as space and so locative case marking also comes to mark time**. Supporting evidence for this comes from data in many Formosan languages where stative verbs or inchoative verbs are marked with the locative *-an* suffix. (Cf. an emotional state such as love, as conceptualized by English speakers, draws on the notion of containment (*in love*, *out of love*).)[1]

Although LOCATION and OBJECT are distinct ontological categories in the grammar of natural language, there is an intuitive sense in which the two are interpermeable in that OBJECTS must exist in a certain space, and LOCATIONS, like objects, can be quantified and are inconceivable without an existent object being situated in it. It is well known that in many languages a nominal entity that is coded by a body-part noun, like 'stomach' for example, often comes to designate a locative relation that highlights the search domain identified with the nominal entity. Thus in many languages a common noun like *stomach* has become a relational 'in'. Common nouns, locational nouns, and adpositions represent different points along this trajectory. This interpermeability and metonymic relation that holds between **object** and **location** give rise to linguistic changes in which Objects are completely reinterpreted as Locations, as in Kavalan, where PF verb forms have merged into LF forms and both now take the LF *-an* suffix.[2] This interprenetrability also spills over into the case marking system. It is generally agreed that semantic cases, especially local cases, can expand to cover syntactic relations such as direct object (Blake 2001). There is evidence that adpositions and case suffixes frequently extend to cover more abstract meanings and to mark syntactic categories. In many Austronesian languages, definite non-subject **patients** (including personal pronouns) may be or must be in the **locative** case form, especially

---

[1] Some examples from Saisiyat should suffice to illustrate the point here:
'ol'ola'-an 'small'; haba'-an 'many'; hehemi'-an 'silent'; hihimi-an 'shy'; inakisa'-an 'like that'; ititi'-an 'a little'; ka-kayzaeh-an 'good'; pakpakaalo'-an 'to be wretched'; simay'akoy-an 'big and fat'; tik-tikot-an 'nervous'; wa:is-an 'powerful'.

[2] As Ross (pc) pointed out, the merger in this case could equally likely have been aided and abetted by the formal similarity of *-en* and *-an*. The point to note is simply that conceptual contiguity between LOCATION and OBJECT must have played an important role.

when the main verb is interpretable as a contact verb, as seen in the data below.

(2) Saisiyat (Pear story): alternation between accusative case and locative case

　　a. … rom-orol 'aehae' ka … (0.8) ka korkoring …am ima
　　　　 follow.AF one Nom Nom child Aux Asp
　　 papama' ka ka-papama'-an ka …. (0.8) ka-papama'-an noka korkoring
　　 AF.ride Acc vehicle Acc vehicle Gen child
　　 'A child who is riding a bike follows.'

　　b. … (0.9) m-wa:i' ka korkoring ima … papama' ray
　　　　 come.AF Nom child Asp AF.ride Loc
　　 ka-papama'-an m-wai:i' ila hiza ray … (1.1) eh=
　　 vehicle come.AF arrive here Loc
　　 'Here comes a child who is riding a bike.'

(3) Saisiyat

　　a. ya'o siae' ka korkoring/no korkoring
　　 1S.Nom happy Acc child Dat child
　　 'I feel happy for the child (because he did something good).'

　　b. sia tikot ka / no haboey
　　 3S fear.AF Acc/ Dat fire
　　 'S/he is afraid of fire.'

　　c. ya'o pa-siba:ah ka rayhil hi 'Obay
　　 1S CAU-lend Acc money Acc/ Dat PN
　　 'I lent 'Obay some money.'

(4) Kavalan

　　a. babar ti-utay timaku / timakuan
　　 AF.punch CL-PN 1S.Acc/ Loc
　　 'Utay punched me.'

　　b. bura ti-utay tu kerisiw timaikuan
　　 AF.give CL-PN Obl money 1S.Loc
　　 'Utay gave me money.'

(5) Amis

　　a. mi-nengneng kaku cingraan
　　 AF-see 1S.Nom 3S.Acc/Loc
　　 'I will see him.'

　　b. cimanan paaca kisu tu futing
　　 who.Acc/Loc sell 2S.Nom Acc fish
　　 'Whom do you sell the fish to?'

In Saisiyat the accusative marker for proper nouns *hi* also functions as the dative marker, as in (3c), although *ini'* is the normal and thus the more expected dative marker. In (4a), the patient pronoun of the verb *babar* 'punch' in Kavalan may be in the accusative or the locative—this being a specific instance of the more general syncretism of LOCATION and OBJECT in the language. In (5a), the patient pronoun of the verb *mi-nengneng* 'see' in Amis may be in the accusative or the locative. Note that the use of the locative or the accusative case marker in these languages does not necessarily depend on the degree of affectedness or different conceptualizations of the force dynamic structure of the event, but is to be interpreted as an instantiation of the mutual permeability of Object and Location.

## 3. Conceptual contiguity of LOCATION and OBJECT

Because of conceptual contiguity, **Locations** sometimes get reinterpreted as **objects**, giving rise to the situation in Saisiyat where PF verbs have encroached upon the domain of erstwhile LF verbs in main clauses. Locations are sometimes mixed up with Objects, giving rise to occasional free variation or at least indeterminacies between the forms in LF and those in PF, as seen in Tsou and Seediq data (6)-(7) below, or even between PF verbs and RF verbs, as seen in the Squliq data (8) below.

Tsou shows some indeterminacy between their PF and LF forms. Native speakers find the following PF and LF forms of verbs interchangeable in meaning:

(6) Tsou
   a. sUcU'Uhi (LF)/sUcU'Uha (PF) 'arrive at', though for its synonym ac'UhU 'arrive at', only the PF form is acceptable: ac'Uha (PF)/*ac'Uhi (LF)
   b. e'hongi (LF)/e'honga (PF) 'smell'
   c. A small number of spontaneous motion verbs must take PF forms to express the idea of motion toward a goal for a purpose, rather than the more expected LF forms: mimio (AF)/mimia (PF) 'to move around for a purpose'; emo'unu (AF)/emo'una (PF) 'go to (a location for a purpose)'

The Seediq pear story narrator in (7) shifts from *psa-un* (PF) 'to put' to *psa-an* (LF) 'to put' in successive utterances:

(7) Seediq

    ma    lnamu    na    we, ga    na    psa-un    dheran    di
    Conj  Past.pick Gen   Top Aux   Gen   put-PF    ground    Part
    … ga    na    psa-an    rawa    truma    ha
       Aux   Gen   put-LF    basket  inside   Part
    'And what he picked was set on the ground … it was put inside the basket.'

In Squliq, sometimes an RF verb may be replaced by a PF verb without any change in meaning, since the RF marker *s-* and some PR verbs are known to be capable of forming causative verbs such as the main verbs in (8).

(8) Seediq

    a.  s-n-galu'            ciwas    qu'    yumin
        RF-Pfv-sympathize    PN       Nom    PN
        'Ciwas sympathized with Yumin.'
    b.  g-n-alu'             ciwas    qu'    yumin
        sympathize-PF.Pfv    PN       Nom    PN
        'Ciwas sympathized with Yumin.'

Interestingly, the LF verbs in Cebuano and Tagalog regularly take the form *gi-…-an* and *-in-…-an* respectively, where *gi-* and *-in-* are past tense affixes for PF verbs and *-an* the locative focus suffix, suggesting that there is an intimate association between LOCATIONs and OBJECTs in these languages, exactly as in Formosan languages. In Taglog, some verbs in PF indicate perfective or holistic, and those in LF indicate imperfective or partitive events, as in (9b):

(9) Tagalog (Aldridge 2002)

    a.  k-in-ain        ko          ang    isda
        eat-PF.Past     lS.Gen      Top    fish
        'I ate the fish.'
    b.  k-in-ain-an     ko          ang    isda
        eat-past-LF     lS.Gen      Top    fish
        'I ate at the fish.' (but I did not finish it.)

As observed in Aldridge (2002), (9a), where the main verb is in PF form, expresses a perfective event, while (9b), where the verb is in LF form, expresses an imperfective and partitive event—the agent did the action of eating, but he did not finish it, or he left some. This strategy of using LF clauses to indicate partitive events is also found in Cebuano, as in (10).

(10) Cebuano

    a.   gi-kaon    ni    Maria  ang    mansanas
        Past-eat    Gen   PN      Nom   apple
        'Maria ate the apple.'

    b.   gi-kan-an  ni    Juan    ang    salad
        Past-eat-LF Gen   PN      Nom  salad
        'Juan ate the salad.' (but he left some/did not finish it.)

Returning to the types of verbs that must take an LF focus marker, the following types of verbs represent an exhaustive list culled from the NTU Tsou corpus:

(11)   Tsou: verb types that take an LF focus marker *-i*

    a.   perception/cognition verbs: aiti 'see'; (cf. iui'ia (PF) 'to watch closely') humi 'see (the footprint of animals)'; cohivi 'know'; talUi 'hear; understand'; e'hongi 'smell'

    b.   motion verbs: mi'usni 'face someone/some location; to address'; sUlUi 'bump into'; eupevai 'depart (from each other)'; ec'Uvi 'pick up'; sUcU'Uhi 'arrive at'; pkaaki 'to leave (someone or some location)'; teelUi 'catch up to (someone)'; nopaezumi 'go through (suffering)'

    c.   contact verbs: toai 'pick (fruit)'; taeoci 'write in advance' (the PF form *taeoca is impossible); e'tUi 'pull'; ena'vi 'to pull at'; pas'osi 'hit; point to'; paceicei 'lick (wound)'

    d.   transaction verbs: fii 'give'; pahsUsUfti 'teach'; tuocosi 'ask'

    e.   goal-directed activity verbs: to'usni 'help'; teo-epUngi 'finish doing'; kinkini 'to weigh'; pnafi 'make sure'; ena'vi 'to pull (at) for a long time'; eahioai 'to work (at)'; nochiei 'take care of'

Turning to Squliq Atayal, verb types that must take an LF focus marker as found in the NTU corpus are the following:

(12)   Squliq Atayal: verb types that take an LF focus marker *-an*

    a.   perception verbs/cognition verbs: p-tnga-n 'peek at' (no PF form); kta-n 'see' (PF form kt-un rarely used); sok-an 'smell' (no PF form); ngay-an 'look at' (no PF form); 'lw-an 'find' (no PF form); pung-an 'hear'; s-in-nhy-an 'believe'

    b.   motion verbs/locative verbs: sy-an 'put' vs. sy-un; qy-an 'hang'; p-hkny-an 'walk (road)'; wah-an 'come to'; kyup-an 'enter'; sa-n 'go to'

      c.   goal-directed activity verbs: sgay-an 'bid farewell to' (realis); khg-an 'take care'

      d.   contact verbs: tsbeŋ-an 'grasp' (no PF form)

      e.   transaction verbs: bil-an 'buy'; tbil-an 'sell'; ksiug-an 'borrow'; pqut-an 'ask' (PF form pqut-un rarely used)

We have shown that the distinction between verbs in PF and in LF forms correlates well with the distinction between change-of-state vs. surface contact verbs and that the nominative NP of a PF clause is used to encode an Object and that of a LF clause a Location. The distinction can still be observed in languages such as Tsou and Squliq, where these two types of focus construction are still vibrant, but shows signs of merger in Saisiyat and especially in Kavalan, given the conceptual contiguity and thus permeability of Objects and Locations. In the next section we turn to the functions of RF clauses.

## 4. Functions of RF verbs

Turning to RF verbs, it is important to appreciate that a great majority of the O arguments in RF clauses can be shown to encode transported theme, an entity that undergoes movement, physically or metaphorically, from one locale to another. These verbs are thus *bona fide* causative motion verbs. In Tagalog, a language which has lost most of the functions associated with RF clauses, for example, there were a total of just two RF clauses in our (as yet very limited) corpus and these are clearly causative motion verbs that take a nominative NP that encodes transported theme: *i-ni-hulog* 'RF-past-fall; throw down' and *i-b-in-igay* 'RF-give-past'. In Tsou, a total of 30 RF clauses representing 20 distinct verb types were found in the corpus. Of these 20 verb types, a vast majority of them (14 out of 20, or 70%) also take a co-occurring locative phrase. This distributional phenomenon can be given a quite natural account in terms of transported theme. These 20 verb types are given below in (11), where parentheses indicate the O arguments the verbs took in the corpus data. Note that what is analyzed here as encoding the transported theme associated with the nominative NPs of these RF verbs has been analyzed by most researchers in the literature as encoding the thematic role of Instrument. It is fairly easy to demonstrate that the **Instrument** analysis is to be dispreferred, and ultimately rejected, since it cannot explain why these RF verbs often take a locative phrase complement, whereas the transported theme analysis can. Most of the RF verbs in these languages simply function to indicate the movement of an entity from one locale to another, with the goal of movement indicated by a locative phrase complement. The localist schema for RF verbs can be represented as follows:

(13)   [Agent] CAUSES [Theme   at   Location]

where the theme is coded as the nominative marker, and the agent as the genitive and the CAUSE predicate corresponds to the main RF verb.

Of direct relevance to the present analysis is the important observation by Ross (1995:756-760) that "There is not sufficient representation of the main clause use of the instrument voice to reconstruct this function of the morphology of Proto Austronesian. It is found widely in nominalizations, however…" Evidence from Tsou, Saisiyat and especially Kavalan supports that observation. Example (14) lists RF verbs whose nominative NPs encode the movement of a transported theme in the Tsou corpus. NPs within parentheses are actual nominative phrases found in the corpus.

(14)   Tsou: RF verbs that encode the movement of a transported theme:
eskoveineni 'make (vine) into a circle'; lehtothomneni 'to wrestle with sb.'; eevozneni 'to hurl (the child) to'; to'seni 'to toss (him) to his child'; eoveineni 'return (hat) to sb.'; faeni 'give (fruit to a child)'; p'aeni 'to feed (children) something'; eupteUlUneni '(two people) meet each other'; p'ateni 'show sb. what he's doing'; poa-cohivneni 'cause sb. to know (my suffering)'; tuop'op'eni 'turn over (a bottle)'; tohphuceni 'release (head) from a bottle'; hooyoeveni 'insert (head) into (the bottle)'; tokUneni 'throw (frog) away'

The remaining RF verbs in the corpus are interpretable as encoding the thematic role of benefactive (i.e. affectee), especially when the O argument is an animate NP. A moment's reflection should convince us that the multiple functions of the nominative NP in RF clauses probably develop along the following lines: a transported theme, acting as it does as the starting point of a motion event, is first interpreted through metaphorical inference to implicate **causation** for certain events (as if a causative motion event is analogized as a metaphorical causal event, with the causee taken by some causer agent to some location in a force-dynamic fashion), and, the **causation** for events, through further inference as a result of conceptual contiguity between what triggers causal events and what is affected, especially in emotion domains, as an **affectee** of those events if the O argument is animate. These causal/affectee NPs are given in (15). The noun phrases in parentheses are again the actual nominative NPs of the RF clauses found in the corpus.

(15)  Tsou RF verbs that encode the thematic role of benefactive (cause/affectee):
pa-teoteaiveni 'cause to do st. (for other people)'; itotonveni 'strive for (the language of Tsou)'; titheni 'spend for (children)'; ananav'eni 'work hard for (the language of Tsou)'; aveoveoneni 'happy about st.'; pa-tuocosneni 'cause to ask about (your opinion on Tsou for sb)'

Whether an NP is to be interpreted as denoting a cause or an affectee is a function of both the semantics of the verb and the nominative NP in question. For example, in {*Pasuya is happy about something.*}, the *something* functions as a cause triggering the emotion of happiness. In {*Pasuya ran for Mo'o.*}, the default meaning is {'Pasuya ran for the benefit of Mo'o.'}, and can cover the following functions: {'Mo'o ran for the sake of Mo'o.'}; or, {'Pasuya ran because Mo'o told him to.'}. In these cases, *Mo'o* can be both a cause and a beneficiary.

Exactly the same mode of analysis can be applied to the RF verbs in Saisiyat. A total of 42 RF clauses were found in the corpus, representing 19 distinct verb types. Of these, a total of 12 verb types (63.1%) take a nominative NP that encodes transported theme. These 12 verbs whose nominative NPs encode transported theme are given below in (l6), and those that encode cause/affectee are given in (17). Note that the noun phrases within the first set of parentheses in (16) indicate the nominative NPs that the verbs take in the corpus data, and those in the second set of parentheses indicate co-occurring locative complements found in the corpus data.

(16)  Saisiyat: 12 RF verbs in Saisiyat whose nominative NPs encode transported theme
si-mari 'take (frog) to some location'; si-'osa 'throw (child) (down the cliff)'; si-tomok 'to insert (head) into (bottle)'; si-paka:I 'hang (child) on the (deer head)'; si-pan'oe'ong 'turn (bottle) upside down'; si-kosiza 'insert (head) into (bottle)'; si-sa:ah 'spill (fruits) (all over the ground)'; si-bay 'give (fruit)'; si-karma 'steal (a basket of fruits)'; si-patono' 'to bump into (rock)'; si-roton 'gather (meat)'; si-sapih 'exchange (bottle for wine)'

(17)  Saisiyat: 7 RF verbs whose nominative NPs encode the semantic role of cause/affectee
si-'oe'oe 'cause (someone or something to be called)'; si-panra:an 'affected by walking (on the stony road)'; sik-ra:iw 'get left behind'; si-ngowip 'affected by forgetting (his hat)'; si-patabil 'cause to be worshipped'; si-wa:i' 'affected by coming'; si-pahila 'affected by being under the sun'; si-paplangi 'affected by sharing (food)'

## **5**. **Tsou: A (nearly) perfect split O language**

A perfect split O language is one in which a verb can occur only in one non-agent focus form, and the appearance of this particular non-agent focus form can be inferred largely on lexical semantic grounds. Moreover, if a verb appears in two non-agent focus forms, then one of which must be in RF/BF, and the nominative NP of that RF/BF verb can only encode the semantic role of benefactive rather than transported theme. In other words, a perfect split O language is one in which the mapping between focus form of a verb and its meaning is predictable. I shall show below, based on corpus evidence, that, allowing for certain intractable data and indeterminacies between PF and LF forms alluded to above in §2, Tsou fits the definition above remarkably well and can thus be said to be a perfect split O language.

Compared with other Formosan languages, the predicate structure of Tsou is unique. Unlike Atayal and Saisiyat, for example, Tsou employs a separate system of auxiliaries to mark TAM, freeing the verbs to mark purely lexico-semantic information. This is why an investigation of the way the focus form of a verb and its meaning is mapped in Tsou should shed light on how O arguments in Formosan languages in general are split.

A preliminary analysis of the data in Tsou yields the results shown in Table 2 below. Table 3 identifies the thematic roles associated with the nominative NPs of the various verb types. (See Huang & Huang (forthcoming) for further detail).

**Table 2: Verb types and acceptable focus forms in Tsou**

| Verb type | AF | PF | LF | BF |
|---|---|---|---|---|
| (1) Activity (i) | ∨ | ∨ | (∨) | ∨ |
| (2) Activity (ii) | ∨ | -- | ∨ | ∨ |
| (3) Placement | ∨ | ∨ | ∨ | ∨ |
| (4) Ditransitive | ∨ | -- | ∨ | ∨ |
| (5) Emotion (i) | ∨ | ∨ | -- | (∨) |
| (6) Utterance | ∨ | ∨ | -- | (∨) |
| (7) Perception and cognition | ∨ | -- | ∨ | -- |
| (8) Motion (i) | ∨ | -- | ∨ | (∨) |
| (9) Motion (ii) | ∨ | ∨ | -- | -- |
| (10) Collective action | ∨ | -- | -- | ∨ |
| (11) Emotion (ii) | ∨ | -- | -- | ∨ |

**Table 3: Verb types and the thematic roles of nominative NPs**

| | Nom NP in PF clause | Nom NP in LF clause | Nom NP in BF clause | Examples PF/BF, LF/BF, PF/LF/BF |
|---|---|---|---|---|
| (1) Activity (i) | Patient | (Location) | Benefactee/ Instrument/ Transported theme | *aUt'Uca/aUtucneni* 'raise; take care of', *opcoca/opcocneni* 'kill' |
| (2) Activity (ii) | -- | Goal | Beneficiary | *p'ecihi/p'ecihneni* 'step on', *tvici/tvicineni* 'weed' |
| (3) Placement | Patient | Source/goal | Beneficiary/ Instruments | *sia/sii/siineni* 'put', *teapha/teaphi/teaphineni* 'fill' |
| (4) Ditransitive | -- | Recipient | Transported theme | *pii/fiineni* 'give', *pa'cohivi/pa'cohivneni* 'teach' |
| (5) Emotion (i) | Stimulus | -- | (Stimulus) | *coveoza/coveozneni* 'feel pity for', *ta'kuv'a/ta'kuv'eni* 'worry about' |
| (6) Utterance | Utterance | -- | (Benefactee) | *eainca* 'say', *pasunaenva, psunaenveni* 'sing' |
| (7) Perception and cognition | -- | Percept | -- | *aiti* 'see', *e'hongi* 'smell' |
| (8) Motion (i) | -- | Source/goal /ground | (Source/goal) | *sUc'Uhi* 'arrive at', *capi* 'climb' |
| (9) Motion (ii) | Ground | (Ground) | (Ground) | *sucaefiza/sucaefizi/sucaefz-neni* 'pass by', *mooveia* 'return' |
| (10) Collective action | -- | -- | Object of sociative Action | *eupteuluneni* 'meet with', *noteuyununeni* 'be with' |
| (11) Emotion (ii) | -- | -- | Stimulus | *kaebneni* 'happy for', *congoneni* 'pain for' |

It is of some theoretical interest to observe that in verb types (5), (8), and (9), the nominative NPs of the BF verbs are used to encode location/goal and thus have encroached upon the traditional domain of LF clauses. Note that it is also in BF verbs where native speakers frequently disagree over their correct usage or the specific forms they take, suggesting that the grammar of BF verbs is undergoing fundamental linguistic changes.

# 6. Excursus: how Cebuano splits its Os

Cebuano, like Tagalog, has lost most of the functions associated with RF clauses in Formosan languages and the only function left now is for their nominative NPs to encode transported theme. The usual RF functions of indicating instrument, cause, or benefactive have now been taken over by LF clauses, as seen in the following examples (provided by Michael Tanankingsing). LF clauses in Cebuao still retain their normal functions of indicating location, goal, and, more generally, imperfectivity/partitive reading.

(18)  Cebuano
    a.  Instrument/Medium

| gi-imn-an | sa | bata | ang | baso |
|---|---|---|---|---|
| Past-drink-LF | Obl | child | Nom | glass |

        'The child frank from the glass.'

    b.  Benefactive

| gi-himo-an=nila | si | nanay | ug | balay |
|---|---|---|---|---|
| Past-make-LF=3S.Gen | Nom | mother | Obl | house |

        'They built a house for Mother.'

    c.  Goal

| gi-sak-an | sa | mga | tulisan | ang | ilang | balay |
|---|---|---|---|---|---|---|
| Past-go.up-LF | Obl | PL | robber | Nom | 3PL.Poss | house |

        'The robbers went up to their house.' (They were robbed.)

    d.  Imperfective/partitive

| gi-kan-an | ni | Juan | ang | salad |
|---|---|---|---|---|
| Past-eat-LF | Gen | PN | Nom | salad |

        'Juan ate at the salad.' (but he left some/did not finish it.)

Instrument/medium is interpretable as a variation on location. If someone flies in the air, 'the air' is a medium, and thus also a location. The beneficiary in (18b) is the intended recipient of the activity of house-building and recipients are typically encoded as the nominative NP of LF clauses.

The Cebuano data seem to suggest that two types of applicative construction can be distinguished: (1) locative and (2) benefactive and instrument, and that (2) seems to depend on (1). In other words, (1) serves as the anchor for the development of (2). However, instrument and benefactive have also been shown to derive from the semantic role of transported theme (see §4). Does this mean different languages may arrive at their RF constructions via different routes, and thus the way RF constructions decline and finally lose out may also differ from language to language?

## 7. Referential emotions in Saisiyat

Different pragmatically driven meanings can become crystallized in different language communities. We have shown in the preceding paragraphs that the nominative NP in the RF *si*-construction in Saisiyat is either a transported theme, or a cause triggering certain events, or an affectee caused by the action performed by the agent marked by the genitive case. It appears that the affect interpretation is the most pervasive element in the *si*-construction in Saisiyat. This means that the nominative NPs of the RF verbs given in (16) are also (emotionally) affected, not just those in (17), unless they are inanimate objects. This reading of affect becomes more salient in situations where the verb does not and cannot involve moving any transported theme, thus barring its NP from encoding transported theme, as in (19), where the child, the unmarked nominative NP, is affected by 'Obay's or their going away.

(19) a. si-osa ni 'Obay, korkoring homangih
RF-go Gen PN child cry.AF
'The child cried (because) 'Obay went away.'
b. sik-a-osa nisia, korkoring homangih
RF-?-go 3.Gen child cry.AF
'The child cried (because) they went away.'

Saisiyat makes an interesting distinction between two types of RF construction. In the *sik-* construction the agent or the speaker must directly perceive the event associated with the verb, while no such precondition is required of the *si*-construction. In (19a), for example, the nominative NP (the child), or the speaker must have seen 'Obay going away. What these two types of construction have in common is that the referent in the nominative is invariably affected. In what follows, we shall term these affect-inducing verbs referential emotion verbs (see Huang 2004 for further discussion).

## 8. Peculiarities of the RF verbs in Squliq Atayal

We now turn to take a look at RF clauses of a third language Squliq Atayal. A total of 61 RF clauses in the Squliq corpus represent 35 distinct verb types. As in Tsou and Saisiyat, a majority of them (18 for 51.4%) take a nominative NP that encodes transported theme, as shown in (20).

(20) **RF verbs in Squliq** whose nominative NPs encode transported theme
s-paNa 'carry (packs of rice)'; s-buli' 'throw away (dowry)'; s-ruruw 'push (door)'; s-bul 'dip (cotton) into water'; s-pqzyu' 'transmit; tell (things)'; s-qni' 'cover up (dream)'; s-hgu 'dip (head) into water'; s-tmaq' 'mash (tobacco leaves)'; s-tmumu 'tie (hairs) into a knot'; s-qihul 'force (boys) to'; s-p-gluw 'cause (omen birds) to follow'; s-biq 'give (chicken)'; s-hoq 'insert (hoe)'; s-'pux 'press (pigeon peas)'; s-'atuk 'make vine into a circle'; s-p-qaniq 'feed (child)'

Note that since these causative motion verbs are signaled by the prefix *s-* , the prefix has come to be reanalyzed as a causative morpheme. This *s-* encodes causatives that express more indirect and less manipulative causation and means something like 'for reason that; for the sake of'. This causative marker is to be distinguished from another marker *pa-* (*p-* in Squliq), which is more direct and more manipulative. An added complexity in the grammar of causatives in Squliq is that *s-* and *p-* can combine to create the causative sequence *s-p-* to express the idea 'to cause something to happen for reason that; to cause something to happen for the sake of'. In this sequence, *s-* is a causative marker, rather than an RF maker for argument-enhancing construction.

The majority of the remaining RF verbs encode transitive (but not causative) events that take a direct object marked by nominative case and in addition also require the presence of another clause somewhere in the prior discourse specifying the cause for the transitive event; (21) is an example.

(21) yaqih qu' tuqiy qa; s-htuy-mu' qu' laqi'-mu' ka
Bad Nom road Det RF-forbid-1S.Gen Nom child-1S.Gen Lig
Musa' m-Niq sa BuN
Go.AF AF-swim Loc river
'Because the road condition was bad, I forbade my child to go swimming in the river.'

These cause-specifying verbs in the corpus are given in (22).

(22) Cause-specifying RF verbs in Squliq
skita' 'see (the omen bird) (because of st.)'; s-kut 'kill (sb.) (in order to---)'; s-htuy 'forbid bad men (to do st.)'; s-kal 'to say (things about the elders) (in order to do st.)'; s-sulin 'burn (stone) (in order to do st.)'; s-qbuyaN 'ambush (sb) (in order to do st.)'; s-pge 'run away to (the mountains) (in order to do st.)'; s-'agal 'use (the vine) (to do st.)'; s-naga' 'wait for (the child) (in order to do st.)'

In both Saisiyat and Squliq the nominative NP encodes the idea of cause, but only in Saisiyat is affect part of the meaning of the RF construction. Referential emotions in Squliq are emotions expressed by verbs that take the RF prefix *s-* and that incorporate the meaning of 'emotional benefactive (or malfactive)', 'emoting because of someone' and 'emoting with someone', as illustrated below (examples taken from Maya Yu-ting Yeh 2002):

(23)  Happiness
      scagaw 'to be proud of'; s-bolas 'to lust for/after'
      a.   s-bolas  yal  ni'  yumin qu'   yuli'  qa
           RF-lust  very Gen  PN    Nom PN    Det
           'Yumin lusts after Yuli.'
      b.   s-cagaw    misu'
           RF-proud   1SG+2SN
           'I am proud of you.'

(24)  Sadness
      sŋloŋ 'to have regard for sb. feeling'; sgalu' 'to sympathize with'; stltu' 'to feel sad for'
      a.   s-ŋloŋ              ciwas    yumin
           RF-have regard for  PN       PN
           'Ciwas has regard for Yumin's feeling.'
      b.   s-tltu'               misu'
           RF-heartbroken        1SG+2SN
           'I feel heartbroken because of you.'

(25)  Fear
      snkux 'doubt for'
      a.   s-nkux    hakaw  qu'   yumin
           RF-doubt  PN     Nom   PN
           'Hakaw is doubtful about Yumin.'

(26)  Anger
      stluhin 'to be angry about'; stlequn 'to be mad about'
      a.   nanu yaw  qani  qu',  s-tlequn ni'  hakaw  qa
           what thing Det  Nom  RF-mad  Gen  PN     Det
           'What is Hakaw mad about?'

(27)  Disgust
      s-sayux 'to be ashamed of'; s-yubiŋ 'to be stingy with'
      a.   s-yubiŋ  yal  ni'  yumin qu'   qsyaq
           RF-stingy very Gen PN    Nom   waer
           'Yumin is very stingy with water.'

In each of the sentences above, the NP in nominative is traditionally taken to be the emotional beneficiary, but is semantically the driving stimulus that causes the NP in the genitive to experience the emotion in question. Seen in this light, it is quite natural to perceive an intimate connection between the nominative NP in causative motion verbs that denotes the locative source of motion (e.g. the verb *carry* and its object noun in (20)) and the nominative NP in referential emotion verbs that provides the stimulus for the emotion experienced by the NP argument in the genitive case (e.g. the adjective *proud* and its oblique object noun phrase in (23b)). In other words, the causative motion meaning serves as the anchor for the development of instrument/benefactive meaning of RF clauses. If one takes the perspective of the experiencer NP coded in the genitive case, then the referential emotion s/he experiences is an emotional effect; on the other hand, if one takes the perspective of the stimulus NP in the nominative, then s/he can be both the emotional beneficiary and the stimulus (or the 'instrument' in the Fillmorean case grammar sense) that triggers the experience of the said emotion, hence the 'instrumental' reading associated with a sentence like (28), where the nominative NP is an inanimate object (*kahoey* 'tree').
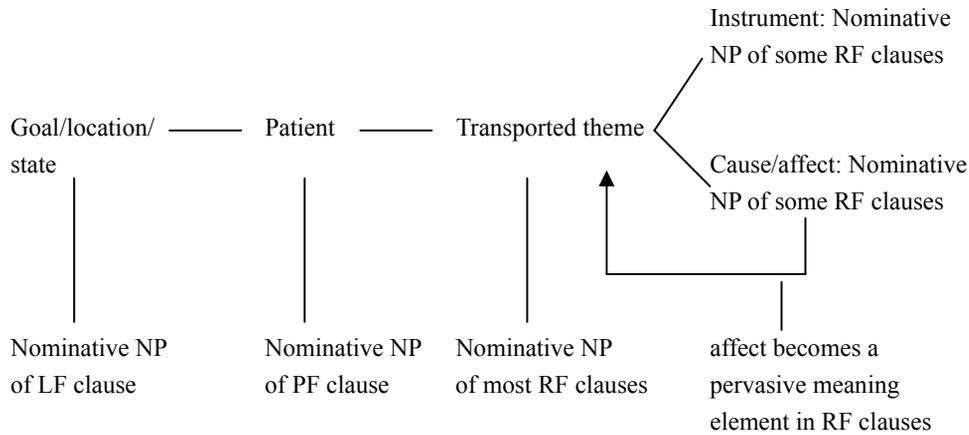
(28)  kahoey  si-Sebet  ni  'oya'  ka  ahoe'
      tree    RF-bet    Gen mother Acc dog
      'Mother beat the dog with a tree stick.'

## 9. Semantic space for O arguments

The O arguments in the applicative constructions, LF and RF clauses, have been shown to encode spatial notions of Goal/Location and Transported theme respectively. Other semantic roles of the O arguments such as Instrument, Benefactive, Cause and the general meaning of affect are shown to be secondary to and derivable from the more basic spatial notions via pragmatic inferencing. It seems quite natural to assume that these spatial notions are psychologically basic and serve as the anchors for the expression of other more abstract categories. A number of researchers (e.g. Anderson 1971, Lyons 1979) have frequently observed that all thematic roles derive from a small number of universal spatial relations. The split O phenomena shown in the preceding sections show that these different types of O can be given a fairly straightforward interpretation.

If we put together all of the coding patterns for the O arguments, we have the semantic space for the O arguments shown in the figure below. (See Croft 2003 for the concept of semantic space.)

Semantic space for O arguments

Instrument: Nominative
NP of some RF clauses

Goal/location/ ———— Patient ———— Transported theme <

state

Cause/affect: Nominative
NP of some RF clauses

Nominative NP                Nominative NP              Nominative NP              affect becomes a
of LF clause                 of PF clause               of most RF clauses         pervasive meaning
                                                                                   element in RF clauses

Note that each 'point' in semantic space is best seen as representing a complex primitive unit, namely the semantic frame evoked by a particular use of a construction. The representation of semantic space is governed by the semantic map connectivity hypothesis (Croft 2003:96), which requires mapping onto a connected region in semantic space. Thus the semantic space proposed for O arguments shown above constrains possible coding patterns for language-specific constructions and thus allows for a cross-linguistic comparison of those constructions. Specifically, it predicts that both the Instrument and the Benefactive/Affect functions of RF clauses are derived from the more basic function of encoding transported theme. This means that once a language has lost some of the functions of RF clauses, it will most likely lose these derived functions first, rather than the other way around, as seen in Tagalog and possibly also in Pazeh (cf. Li & Tsuchida 2002).

The semantic space hypothesis also allows for the permeability of various coding functions. In Kavalan, PF verb forms have merged into LF forms and now both take the -*an* suffix. In Thao there is also a partial merger of this kind (L. Huang 2000). The hypothesis would also predict that there is likely to be merger between RF and PF in either form or function, but less likely between the more distant LF and RF forms or functions. Further research will be needed to determine if these predictions are borne out by the empirical data.

## 10. Conclusion

Viewing the focus constructions as applicative constructions has the virtue of allowing us to see that they in fact exhibit a split O phenomenon, and to predict that

languages with applicative constructions in general do not have syntactic processes like passive or antipassive that create further oblique arguments. The search for simplicity is perhaps the most basic theme of all science. As the late Nobel Laureate Herbert Simon put it somewhere, the purpose of science is 'to find meaningful simplicity in the midst of disorderly complexity'. We have shown that all of the thematic roles of the nominative NPs in PF, LF, and RF clauses encode basically some kind of spatial relation. The nominative NP of a PF clause is a Patient object, that of the LF clause an abstract Location, and that of the RF a Transported theme. Other functions of RF clauses have been shown to derive from this more basic function.

The Localist interpretation of the LF vs. PF vs. RF constructions yields an analysis of the O arguments as a split phenomenon and provides a unified and substantive explanation for the forms and functions of the LF, PF, and RF clauses and related grammatical constructions. A number of predictions made about the interaction of these clauses have been pointed out. From a cognitive science perspective, these results show how rich language data, if properly understood, can serve to reveal aspects of high level mental representation.

Applicatives in Formosan languages are not just transitivizing constructions. They are also often called upon to form morphological causatives beyond those formed by attaching the regular causative marker *pa-* to various types of verb stems. These twin strategies suggest that Formosan languages belong to what Nichols et al. (2004) have termed transitivizing languages, rather than de-transitivizing languages, which in turn gives rise to the conjecture that Formosan languages appear to have lexical causatives with much less frequency than a non-applicative language such as English. These are issues that merit further investigation, as is the interesting issue of the interaction between applicatives and causatives.

Shuanfan Huang

# References

Aldridge, Edith. 2002. Nominalization and *wh*-movement in Seediq and Tagalog. *Language and Linguistics* 3.2:393-426.

Anderson, John. 1971. *The Grammar of Case: Toward a Localistic Theory*. Cambridge: Cambridge University Press.

Blake, Barry J. 2001. *Case*. Cambridge: Cambridge University Press.

Blust, Robert A. 2003. *Thao Dictionary*. Language and Linguistics Monograph Series Number A5. Taipei: Institute of Linguistics, Academia Sinica.

Comrie, Bernard. 1978. Ergativity. *Syntactic Typology: Studies in the Phenomenology of Language*, ed. by Winfred Lehmann, 329-94. Austin: University of Texas Press.

Croft, William. 2003. *Typology and Universals*. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1979. Ergativity. *Language* 55.1:59-138.

Egerod, Søren. 1999. *Atayal-English Dictionary* (2nd edition). Copenhagen: The Royal Academy of Sciences.

Ferrell, Raleigh J. 1972. Verb systems in Formosan languages. *Langues et techniques*, ed. by Jacqueline M. C. Thomas and Lucien Bernot, 121-128. Paris: Klincksieck.

Huang, Huei-ju, and Shuanfan Huang. (forthcoming). Verb semantic, focus type and split O in Tsou. Paper to be presented at the Tenth International Conference on Austronesian Languages. Palauan, Philippines.

Huang, Lillian M. 1995. *A Study of Mayrinax Syntax*. Taipei: Crane.

Huang, Lillian M. 2000. *A Thao Reference Grammar*. Taipei: Yuanliu.

Huang, Shuanfan. 2004. Emotion and affect in the syntax of Saisiyat. Paper presented at Syntax of the World's Languages. Leipzig, Germany.

Huang, Shuanfan, and Huei-ju Huang. 2003. On the status of reality marking in Tsou. *Taiwan Journal of Linguistics* 1.2:1-34.

Li, Paul Jen-kuei, and Shigeru Tsuchida. 2001. *Pazih Dictionary*. Language and Linguistics Monograph Series Number A2. Taipei: Institute of Linguistics, Academia Sinica.

Lyons, John. 1979. Knowledge and truth: a localistic approach. *Function and Context in Linguistic Analysis*, ed. by D. J. Allerton et al., 111-141. Cambridge: Cambridge University Press.

Nichols, Johanna, David Peterson, and J. Barnes. 2004. Transitivizing and detransitivizing languages. *Linguistic Typology* 8.2:149-211.

Ross, Malcolm. 1995. Reconstructing Proto-Austronesian verbal morphology: evidence from Taiwan. *Austronesian Studies Relating to Taiwan*, ed. by Paul Jen-kuei Li et al., 727-791. Taipei: Institute of History and Philology, Academia Sinica.

Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. *The History and Typology of Western Austronesian Voice Systems*, ed. by Fay Wouk and Malcolm Ross, 17-62. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.

Starosta, Stanley. 2001. Austronesian 'focus' as derivation: evidence from nominalization. *Language and Linguistics* 3.2:427-478.

Szakos, Joseph. 1994. *Die Sprache der Cou: Untersuchungen zur Synchronie einer Austronesischen Sprache auf Taiwan*. Bonn: University of Bonn dissertation.

Yeh, Maya Yu-ting. 2002. *Emotion Concepts in Squliq Atayal*. Taipei: National Taiwan University MA thesis.

Yeh, Marie M. 2003. *A Syntactic and Semantic Study of Saisiyat Verbs*. Taipei: National Taiwan Normal University dissertation.

Zeitoun, Elizabeth. 2000. *A Bunun Reference Grammar*. Taipei: Yuanliu.

Graduate Institute of Linguistics
National Taiwan University
1, Sec. 4, Roosevelt Road
Taipei 106, Taiwan
sfhuang@ntu.edu.tw

# 台灣南島語的受詞分裂現象

## 黃宣範

### 國立台灣大學

　　本文主張台灣南島語中之非主事者焦點句可以視爲一種應用句式；其中及物受詞 O 表現出有規則分裂現象，稱爲 split O，有些 O 只能出現在受事點句，有些只能出現在處所焦點句，有些只能出現在參考焦點句。不同作用的 O 構成一個語意空間。語意空間的概念可以解釋並預測不同句構之間的演變與互動關係。

關鍵詞：受詞分裂，焦點，台灣南島語，語意空間