THE OCEANIC SUBSTRATE IN MELANESIAN PIDGIN/CREOLE REVISITED: A TRIBUTE TO ROGER KEESING

GILLIAN SANKOFF

1. INTRODUCTION[1]

Anyone who speaks an Austronesian (An) language located in the geographical area where the contemporary Bislamic languages – Bislama, Pijin 'and Tok Pisin – are spoken, and who has also learned one of these languages, has had the shock of recognition at finding unsuspected parallels and commonalities, especially in the realm of syntax, semantics and discourse structure. As we would say in my own 1970s vintage Tok Pisin: *"Kas! em olsem ol i tok long Tok Ples ia!"*

In 1979 Camden (1979:53, cited in Goulden 1990:26), for example, used the Vanuatu language Tangoan as a basis of comparison with Bislama:

> not because it has any special relationship with Bislama not shared by other...Oceanic type languages...Rather, Tangoan is regarded as broadly representative of this substantial group of Oceanic type languages, with the implication that a comparison between Bislama and any other such language...would probably show comparable results.

In 1985, Simons (1985:54, cited in Goulden 1990:26) explained her use of To'abaita in comparing it with Solomon Islands Pijin (SIP) as follows:

> I do not mean to imply that To'abaita is any more closely related to SIP than any other of the Malaitan dialects...But I do assume that To'abaita is representative of all the Malaitan languages and therefore it is appropriate to use it specifically in making a comparison to SIP.

Goulden himself acknowledges that his knowledge of Lusi led him to spot parallels between Lusi and other West New Britain languages on the one hand, and Tok Pisin – in particular the variety that Goulden (p.6) refers to as "distinctly New Britain" – on the other.

In my own case, the *tok ples* or local language in question is the Austronesian language I learned in 1966-67 when I first went to Papua New Guinea: Buang, a Morobe Province An language of the Central New Guinea branch. In several of my papers on Tok Pisin (TP), I

---

[1] I am indebted to a number of scholars present at FICOL, as well as others with whom I have corresponded or talked in person, for helpful discussion of issues raised in this paper. In particular, I wish to thank John Lynch, Terry Crowley, Ann Chowning, Suzanne Romaine, Nick Faraclas, Andy Pawley, Peter Mühlhäusler, Miriam Meyerhoff and Bill Labov. My own fieldwork with Suzanne Laberge in Lae, 1971, forms the basis for much of the data presented here, and I thank Suzanne as well as the many Lae area residents who helped us and agreed to be recorded for this study. In so far as I have failed to learn from these people or follow their good advice, it is my own fault.

421

have drawn on Buang as an initial inspiration for parallels, and then explored other Melanesian Austronesian (MAn) languages in which similar structures are to be found.

Of course, it was Roger Keesing (1988), who brought most strongly into prominence the idea that what he called Melanesian Pidgin was modelled very closely on the substrate languages of the south-western Pacific. Keesing's proposal was that the languages of the Eastern Oceanic branch of MAn languages, including "the languages spoken in the eastern Carolines, Rotuma, Fiji, and the Gilberts" (p.69) as well as, of course, the languages of the Solomons and Vanuatu, were the most significant in moulding the nineteenth century Pacific pidgin that was antecedent to the modern Bislamic languages. This work of his is an ambitious attempt to delineate "the core syntactic structures of Oceanic languages" (p.69), and to show in breadth and in depth how Melanesian Pidgin offers "unmistakable evidence of the stamp of Oceanic grammar" (p.119). In Keesing's own words (p.106), he

> sought to provide a framework in time and space within which plausible sociolinguistic processes of substrate influence, in 'conspiracy' with superstrate influences and universal constraints and faculties, could have operated...In doing so, I have provided a plausible explanation of why it is possible to take *any Oceanic language of the Southwestern Pacific* and, in comparing it with a dialect of Melanesian Pidgin, make a case for substrate influence [emphasis mine].

Part of Keesing's rich legacy to Pacific linguistics is, then, the forceful case he made for a vision of Solomons Pijin, Bislama and Tok Pisin as Oceanic An languages. He felt that the substrate should take centre stage in our analyses of the Bislamic languages. But if I share many broad areas of agreement with Roger, I also disagree with him on many points, and this makes his absence from among us all the more difficult to bear. I would have loved to spend more time arguing with him. Indeed, I think I will be debating with Roger for many years to come. It is a tribute to him that his contribution was so rich as to provide us with so much meat that it will take the rest of us a long time to chew. It is in this spirit of exchange, of respect, of love of controversy, but above all of the shared joy of uncovering the historical building blocks that shaped the evolution of the modern languages of the south-western Pacific, that I dedicate this paper to Roger Keesing's memory.

Keesing's scenario of the development of the Bislamic languages in the nineteenth century is one controversial aspect of his work, another being his downplaying of variability. His statement (1988:53) that until the 1880s there was a single dialect of Pacific Pidgin would initially seem to reflect the consensus already established by other scholars.[2] Mühlhäusler (1978:81), for example, speaking of the first decade of existence of Samoan Plantation Pidgin (SPP), 1867-79, states that:

> Because of a number of factors, including common recruiting grounds for most Pacific plantations and a number of linguistic conventions that had emerged in Pacific Jargon English, this early form of SPP did not differ greatly from the plantation pidgins found in Queensland or New Caledonia.

---

[2]   It is clear, however, that there are important differences among the accounts provided by Mühlhäusler (1976, 1978), Clark (1979-80), Keesing (1988) and Crowley (1990a). A sociolinguistic analysis of early contacts in the pidginisation process in Queensland appears in Sankoff (1985); my own historical synthesis will appear in a book-length treatment currently in preparation.

However, both Mühlhäusler and Crowley have acknowledged, and dealt head-on with variation at all levels (including that found in the written documents of the nineteenth century),[3] and both reject Keesing's emphasis on the structural uniformity of nineteenth century Pacific Pidgin. Crowley (1990a:40), for example, states that:

> It is almost certain...that at any particular time, there were also structural differences between one variety of Melanesian Pidgin and another, even if these were only small, and even if these only represented greater preferences towards one variant over another.

Mühlhäusler (1985b:36) documents variation and discontinuity in most of the Pacific pidgins, explaining that this discontinuity has been a result of at least four major factors: "rapid changes in population composition and population movements,...different functional requirements, changing patterns of language transmission,...[and] planning or other outside interference". In this paper, as in all of my previous work, I take the analysis of variability to be crucial to analysing the structure of TP. I will not deal further here with the historical issues still being debated and investigated by several scholars including myself, except to make it clear that, as far as TP is concerned, I endorse the general conclusions reached by Mühlhäusler (1976, 1978) and Mosel and Mühlhäusler (1982). It is clear that a separate TP tradition began with the approximately 1,000 workers from the Bismarck Archipelago who served on plantations in German Samoa between 1879 and 1889 (Mosel & Mühlhäusler 1982:167). TP continued its development on the German plantations in the Bismarck Archipelago itself during the 1890s, and after 1900, in both that area and on the mainland of New Guinea (Kaiser Wilhelmsland).

The question of substrate per se is the particular focus of the present paper, in which I will deal mainly with TP, making comparisons with SIP and Bislama where possible. To date, two major monographs have been devoted to this question. Mosel (1980) deals with Tolai alone, focusing mainly on lexical and morphological properties. Goulden (1990) surveys 27 languages from fourteen different subgroupings of the Oceanic branch of An, including 11 from Vanuatu, 7 from the Solomon Islands, and 9 from Papua New Guinea (7 of the 9 being from New Britain and New Ireland). In carrying out meaningful comparisons across the contemporary Bislamic languages, we are fortunate in having a good deal of new research that has appeared since Keesing's work went to press. On Tok Pisin, there is further work by Mühlhäusler (1987, 1990) and myself (Sankoff 1990, 1991, 1993, 1994; Sankoff & Mazzie 1991), as well as contributions by scholars who are newer to the Tok Pisin scene: Romaine (1990, 1993); Faraclas (1988); and Verhaar (1991). On Bislama we have Tryon (1987); on Pijin there is Jourdan (1985a,b) and Simons (1985). Studies on all three languages are found in Verhaar (1990). Most important, however, is Crowley's rich contribution (1990a). Not only has Crowley's research added enormously to our knowledge of the development of the contemporary grammar of Bislama, but it has also provided the most detailed comparisons to date of many crucial grammatical features in all three of the modern Bislamic languages.

In this paper, I discuss substrate impact on Tok Pisin grammar in general, and illustrate with features which have already been debated in the literature, and others which I have

---

3   Keesing (p.41) notes that "almost all observers [in the nineteenthth century] have heavily anglicized their renderings of pidgin" yet, as Crowley points out, Keesing does not draw the inference that this problem casts doubt on his assertion of structural uniformity. Crowley's view is that "one could just as easily argue that what was uniform were only Europeans' *perceptions*" (1990a:40).

discovered in recent research. Where possible, I provide comparisons from the literature on the state of parallel phenomena from the two sister languages. What we can conclude about substrate influence may eventually help us assess some of the important unresolved questions having to do with the history of the Bislamic languages. I deal with features that represent (1) grammatical categories whose underlying semantics are typical of MAn languages; (2) discourse patterns; (3) interclause relations; (4) the development of both pre-verbal and post-verbal tense, mood and aspect marking in TP; and finally (5) the 'core' grammatical pattern as proposed by Keesing involving subject-referencing particles (SRP).

2. GRAMMATICAL CATEGORIES WITH UNDERLYING MAN SEMANTICS

As pointed out recently by Goulden, no reasonable person 'could come to a conclusion other than "substrate influence" to account for grammatical categories like those found in the Tok Pisin pronominal system: the inclusive/exclusive distinction, and the robustness of duals and (more weakly) trials. Another such feature common in rural Tok Pisin at least until the 1970s was the 'same sex' versus 'cross sex' semantics of the terms for siblings: *barata* and *susa* respectively. This resulted in confusion on the part of many an expatriate who assumed that these terms held fixed-gender referents as in English, rather than being tied to the sex of speaker as in many MAn languages. Another such feature is the use of the reciprocal 'other' in constructions where English would use 'the one' versus 'the other' (Keesing (1988:241, 246) discusses this in SIP, and it is also a feature of Tok Pisin). Features that are highly marked, especially characteristic of a substrate Sprachbund, and not attested in relevant superstrates or other creoles, are indeed the easiest to identify as sources for pidgin/creole grammatical patterns. But their very markedness also seems to mark them as likely transmission casualties.

The obvious case in the Bislamic languages would seem to involve the distinction between alienable and inalienable possession, about which Keesing (pp.117-118) says:

> Despite the near universality of the distinction of inalienable possession in Oceanic languages, it is not surprising that this distinction is neutralised in Pacific pidgins...The path of simplification here entails a "developmental conspiracy" in relation to superstrate speakers, to whom the alienable-vs.-inalienable distinction would have been unacceptable or opaqu...In the real scenario, they clearly did not. I infer that this kind of neutralisation of surface distinctions and markings would be a fundamental process in the formation of pidgins even in the limiting case where those who contribute to its formation all speak related languages.

One problem with this line of argumentation is that there are numerous counter examples. If Keesing's statement held in general, TP could not have developed the confusing (to Europeans) kinship semantics of siblings, nor the unnecessary (to Europeans) distinctions that force one to indicate whether the 'plural' is in fact only a couple, or to make explicit whether or not the hearer is included as a referent in the 'first person plural'. More importantly, however, with reference to inalienability in particular, I think that Keesing (and, in all fairness, Goulden and Crowley after him) was too hasty in concluding that the grammatical category of inalienability is not represented in the Bislamic languages.

What, after all, does 'inalienability' mean? It means that an object like a mother, or an arm, or a tail, cannot exist without being possessed. There is no such thing as 'an arm': it

has to be somebody's arm – mine, yours, or a third party's. That the alienability/inalienability distinction is typically grammaticalised in MAn languages has been recognised for at least twenty years (Lynch 1973). Now if you are a speaker who has learned that there is a certain class of objects that is always referred to as being possessed, you are going to tend to refer to them consistently as being possessed, even in a language that does not have a special morphological means of doing so. It is Sapir's notion of "habitual modes of expression" that speakers carry over from their first languages into new languages, and it is these habits that eventually lead to the grammaticalisation of distinctions that speakers consistently make.

Consider the following narrative fragment:[4]

(1) a. *Na,* **liklik barata bilong en** *ia   i-wari tru;*   [Mr G.Z.]
         and little   brother POSS him DEM worry very

     b. *na em i-tok strong na em i-krai.*
         and he   speak loud   and he cry

     c. *Tok, "O, yumi no   ken slip long hul bilong ston ia   nogat.*
         say oh we NEG can sleep PREP hole POSS stone DEM no

     d. *Yumi mas go slip long   arere long – as   bilong diwai ia".*
         we   must go sleep PREP side PREP base of   tree DEM

     e. *Orait,* **bikpela barata bilong en** *tokim* **mama bilong en** *olsem*
         so   big   brother POSS him tell   mother POSS him thus

         nau...
         then

     a. And, **his little brother** was terribly worried;
     b. and he shouted and he cried.
     c. He said, "We mustn't sleep in this cave.
     d. We have to go sleep along side of – at the base of this tree".
     e. So, **his big brother** told **his mother** as follows then...

The free English translation seems quite clear. The story is about the middle brother in a family, whose little brother was terribly worried and whose big brother then said something to their mother. But there is no such person. There are only two brothers, and the usage 'his brother' is always reciprocal. A free English translation that didn't mislead English readers as I have done above would have to phrase it: 'the big brother' and 'the little brother'. This

---

4   Abbreviations used in this paper are as follows:

ADJ	adjective	LOC	locative
AUX	auxiliary	NEG	negative
CLIT	clitic	PART	particle
COMP	completive	PL	plural
CONT	continuative	POSS	possessive
DEI	deictic	PREP	preposition
DEM	demonstrative	PRF	perfect
DET	determiner	PUNCT	punctiliar
DU	dual	RHET	rhetorical
EXC	exclusive	SG	singular
FOC	focus	SRP	subject referencing pronoun
INTENS	intensifier	TRS	transitiviser
INTERJ	interjection		

segment contains three of the 276 noun phrases (NPs) which Claudia Mazzie and I classified as "semantically inalienable for speakers of MNAN [=MAn] languages" in our study of TP determiners (Sankoff & Mazzie 1991). We had undertaken the study in order to see whether there was any evidence that distinctions like specfic/non-specific, or definite/indefinite, showed a tendency towards grammaticalisation in TP, and our total sample included 1,865 NPs from the recorded speech of 16 TP speakers. We looked at many possible influences on whether an NP was determined or occurred bereft of any determiner. 'Inalienability' (according to what we took to be typical Melanesian semantics) was the biggest single factor. Table 1 presents the results, confirming our hypothesis that TP speakers were encoding nouns that Austronesian speakers tend to code as 'inalienable' with possessives at a much higher rate than other nouns. Whereas only slightly more than half (53%) of nouns in general occur with a determiner of some kind, 85% of 'inalienables' do, and overwhelmingly, this determiner is a possessive.

TABLE 1: CONTRAST BETWEEN ALIENABLE AND INALIENABLE NPS IN THE SPEECH OF 16 TP SPEAKERS, IN TERMS OF CO-OCCURRENCE WITH DETERMINERS AND POSSESSIVES

	INALIENABLE NOUNS	ALIENABLE NOUNS	ALL NOUNS
% determined	85%	53%	58%
% possessed	74%	12%	21%
N	(276)	(1589)	(1865)

As Capell (1969:42) remarked, determiners are not a typical grammatical category in the languages of the area, though they do occur in some. Inalienability, however, is a category in language after language, and apparently the speakers of Tok Pisin also find a means to express it, marking nouns as being possessed in contexts where many other languages would use a different category of determiner. Keesing is right in stating that we do not find a morphological distinction between two types of possessed objects; nevertheless Tok Pisin speakers in their habitual modes of expression appear to reflect the infuence of the inalienable category in very consistently giving it expression. Such habits have, in other cases, led to grammaticalisation of distinctions, and the habits of rapid speech here as well might eventually lead to the kind of morphophonological distinctions that separate alienable from inalienable possession in other MAn languages.

For those familiar with An languages, it is surely not surprising to find another way in which one of the modern Bislamic languages encodes a distinction that shows underlying Oceanic semantics. Doubtless more such parallels will be brought to light in future work in the area. In beginning my evaluation of substrate influence by discussing the case of inalienability, however, my major point is not simply to add one more case to the list. Rather, I intend it as an illustration of the possibility of uncovering important tendencies and patterns in language by a systematic analysis of language use – more specifically, by the quantitative analysis of a corpus derived from natural speech. From the point of view of pidgin/creole studies, there are two major advantages to be gained from the use of quantitative methods. The first is the suitability of quantitative methods for dealing with the variability one typically finds in pidgin and creole languages. The second is the appropriateness of quantitative methods for indicating the kinds of trends in speech that reflect grammaticalisation processes, about which more in the following sections.

3. DISCOURSE

In this section I discuss two discourse patterns: the first, clearly related to An languages; the second, possibly related to Papuan languages.

3.1 FOCUS

The first concerns the evolution of an apparently substrate-related focusing strategy in TP as illustrated in (2) and (3).[5] In (2), the best English gloss (one offered by most TP dictionaries) seems to be the reflexive. In (3), however, *yet* focuses on the responsible party, *yu*, and seems here better translated with a cleft, one of the constructions used for focus in English.

(2)   *Ol Pisin na mi   harim.   Ol **yet** harim   tu.*   [Mrs M.T.]
       3PL TP   and 1SG understand 3SG FOC understand too
       They speak Pidgin and I understand. They understand it **themselves** too!

(3)   *Tok "Orait yu   **yet** kilim pikinini bilong mi".*   [Mr Do.]
       say alright 2SG FOC kill child   POSS 1SG
       (She) said, "Alright, **you're the one** who killed my child".

Yet has several other functions in Tok Pisin. Negative and positive temporals are illustrated in (4) and (5) respectively from the speech of Chavi, a Sepik area man who had apparently learned TP on New Britain as an indentured labourer in the 1920s. Interviewed by Margaret Mead, transcriptions of Chavi's texts were published in Hall (1943).

(4)   *Mi no   savi tok **yet**, mi   nufela boi.*   [Chavi, in Hall 1943]
       1SG NEG know talk yet 1SG new   boy
       I didn't know how to talk [Pidgin] yet; I was a new boy.

(5)   *Bel bilong mi   i hot **yet** i stap.*   [Chavi, in Hall 1943]
       belly POSS 1SG hot still stay
       I am still angry [my belly is still hot].

Lastly, it is used as an intensifier:

(6)   *Oloman! mitupela   i-kam longwe **yet**!*   [Mr T.D.]
       INTERJ 1DU.EXC come far   INTENS
       Man! the two of us came from **very** far away!

Several other particles serve as intensifiers in Tok Pisin, with uses quite parallel to those of *yet* in (6). These include *tru* and *moa*, both of which are well attested with *longwe* (*longwe tru*, and *longwe moa*), as well as *stret* (as an intensifier, almost uniquely confined to the child native speakers in our 1971 sample). None of these forms, however, is used to focus on nominals, in a way that would be parallel to *yet* as a focus particle. Table 2 shows the distribution of *yet* in the corpus I examined, according to the various functions I have just discussed. I was astonished when I tabulated these cases to discover that not a single 'noun + *yet*' example occurred. Though a sentence like (7) sounds fine to my ear, such a sentence, with noun + *yet*, is unattested in my data.[6]

(7)   **Lina **yet** i-kam.*

---

5   For further details, see Sankoff (1993).
6   *Wanpela* and *tasol* are to some extent as focus markers/limiters with nouns.

TABLE 2: INSTANCES OF *yet* ACCORDING TO VARIOUS FUNCTIONS FROM THE 1930S TO
THE 1970S [ADAPTED FROM SANKOFF 1993]

	1920s Chavi	1960s Mope	1971 Adults	1971 Children	Total
temporal - neg. 'not yet'	11	8	-	1	20
temporal - pos. 'still'	3	10	8	6	27
intensifier w. adj/adv.	-	3	3	-	6
focus particle w. pronoun	4	8	13	4	29
Total	18	29	24	11	82

To judge from the distribution of Chavi's examples, it looks as if *yet* came into the language as a negative polarity item, likely modelled after the English 'not yet'.[7] 'Not yet' denies that an expected change has happened; 'still' references an unexpected prolongation. Chavi had only a few positive instances of *yet*, as in (5), but these became preponderant later on. Whereas *yet* with negative occurred much more often than with positive for Chavi (11 as opposed to 3); in the 1960s, the positive uses outnumbered the negative (10 as opposed to 8). Negatives have virtually disappeared in the speech of the adults and children recorded in 1971 (1 as opposed to 14). I hypothesised something like the following as a potential path of evolution:

'not yet' ---> 'still' ---> INTENSIFIER ---> FOCUS PARTICLE

The problem was, however, that the 'generalised intensifier' use did not appear with Chavi. How, I wondered, could *yet* have made the transition between a temporal particle with primarily negative uses and a focus particle without passing through the 'generalised intensifier' stage? At this point, I decided to investigate focus in MAn languages, in which I found a number of interesting parallels.

Many MAn languages use postposed affixes or particles for emphasis or focus. Senft (1986:116) describes two 'emphatic' morphemes in Kilivila, *-ga* and *-la*. The many examples he cites are highly reminiscent of Tok Pisin:

(8)   *Yegu-la   ba-lukwe-m   liliu Tudava.*   [Senft 1986:117]
      I-**myself** I.will-tell-you myth *Tudava*
      I myself will tell you the myth of *Tudava* (and no one else).

(9)   *Kai-tala-ga   beku e-sisu wala o gu bwala besatuta.*
      stone-one-**indeed** stone it-exist only in my house now
      blade   blade
      Indeed, there is just one stone-blade left in my house now.

Senft (p. 54) also describes "a fourfold series of emphatic pronouns" (myself, yourself, etc.) in Kilivila, supplying many examples demonstrating meanings glossed both as 'self' and 'just' (cf. TP *tasol*). The fourth set of emphatic pronouns is found in usages Senft describes

---

7   According to the *Oxford English dictionary* (1994), the use of *yet* in the sense described as "implying continuance from a previous time up to and at the present (or some stated) time" is "archaic or dialectal except in negative context". The most recent example they cite is "You know you look ill yet, very ill" (J.S. Winter, *Bootle's children*, v., 1888). About this construction "with negative preceding", the OED says "the more usual, now the only regular, construction". Though clearly the positive construction was still used to some extent in the nineteenth century, it is likely that the negative was the model for early TP speakers (among other things, it would have been heard with emphatic enunciation, and in shorter strings such as 'not yet' in answer to a question).

as 'idiomatic' (e.g. one can answer with the equivalent of *mi yet*, 'I myself', when asked 'How are you?' or 'Who's there?'). Similar 'idiomatic' usages are also found in the several other Austronesian languages I looked at. Kilivila shares with TP not only the idiomatic use of a 'pronoun+ focus marker' construction, but also the formal distinction between the focus markers used with pronominals and those that apply to other parts of speech.

Another MAn language, Buang, shows similar patterns. Buang has one set of postposed emphatic markers specific to pronouns, and marked for person (the first person is *ngau*, so, for example, the response to 'Who did it?' is *ke ngau* (TP *mi yet/tasol*) 'I myself', 'I'm the one'). Buang also has another invariant marker *mu*, used both with nouns and pronouns.

In Manam, a third MAn language, Lichtenberk (1980) describes two 'intensifiers', *-tina* and *-tu?a*, both used with nouns and the former also with verbs, the meanings glossed variously as 'very, real, genuine, exact, just, much'. A 'limiter' (like *tasol*), *-ba(ya)*, is also "used to form the persistive aspect" (p.204). There is also the 'specifier' *-ma*, meaning 'the same, very', as in "That's the very man!" (p.364). Then there are two 'focus markers': the clitics *-?a*, "always the rightmost element of the constituent that carries the focus marker" (p.477), and *-be* (p.480 ff.). Lichtenberk glosses the focused constituents in English with clefts or underlining, noting that *-?a* represents a "stronger focus" than *-be*. He notes (p.481) that *-be* "is especially common...with emphatic pronominal forms, with emphatic numeral expressions...with locative expressions that refer to a point of origin...and with temporal expressions that refer to a time in the past". Again, we see that despite the applicability of each focus particle or intensifier to several parts of speech, there is some specialisation. In both these respects, the situation is parallel to TP. There is also the interesting connection between a form that can be used both as an intensifier and as a marker of persistive aspect, perhaps just the connection that is made between *yet* 'still' and *yet* intensifier and focus marker. Parallels with TP are clear in the overlap between the semantic domains to which the various intensifiers and focus markers in Kilivila and Manam apply.

In Kwaio, the postposed marker *mola*, referred to by Keesing (1985:93) as a "limiting qualifier", occurs with both nouns and verbs. In the former case it usually means 'only', and occurs only when "the head noun is quantified by a numeral" and when "the speaker is emphasizing that the total number of entities is small". In the nominal paradigm, this is apparently the closest *mola* gets to the 'emphatic' function, but in the verbal system, we see a little extension towards the emphatic uses common in the other languages. With verbs, *mola* usually means 'just' or 'only', but Keesing also notes that it "sometimes conveys the meaning of 'really' rather than 'just' or 'only'". Though some of the functions of Kwaio *mola* overlap with the Kilivila, Manam, Buang and TP 'limiters', it may be that the use of such a qualifier as a focus marker is particular to the New Guinea Austronesian languages we have examined. This is consonant with the idea that Eastern Oceanic languages may have set some of the general patterns found in TP as well as in its sister languages, but the closer parallel TP shows with the New Guinea area Austronesian languages is, I believe, the result of the heavy influence it underwent from Tolai.

In Mosel's *Tolai texts* (1977), we find a particle that is remarkably parallel to the forms we have been examining. Not only is it parallel in function, but its form is a happy surprise

to anyone seeking an explanation for TP *yet*: the Tolai equivalent is *iat*. Examples from Mosel's texts (her glosses) [8] include:

Focus marker with a pronoun:

   (10)   *Iau* **iat**   *pa 'au manga nukure mala ra   tinata i*
            I   PART not I   very   know well DET story of

            nam ra   tubuan   [Mosel 1977:85-87]
            DEM DET *tubuan*
            I **myself** don't know very well the story of that *tubuan.*

Intensifier with a temporal:

   (11)   *...nam ra*   *marum* **iat**...   [Mosel 1977: 78, 82]
            DEM   DET   night   PART
            ...that **very** night...

Focus marker with a noun:

   (12)   *Dia   ga pait nam ra   tubuan, a   Tolai* **iat**
            they.4 TA make DEM DET *tubuan* DET Tolai PART

            dia   ga pait ia.   [Mosel 1977: 86, 88]
            they.4 TA make it
            The Tolai **themselves** made this *tubuan,* they made it.

[I think a closer translation might be: 'They made this *tubuan*, the Tolais **themselves** made it'.]

     Mosel (1980) includes *iat* in a list of Tolai words considered to be possible sources for TP lexical items, and glosses it as a 'particle'. The Methodist Overseas Missions dictionary (1964:54) contains more information:

> **iat 1.** emphatic part., myself, yourself, etc.; **ia iat** he himself, **iau iat,** I myself. As such it is also used as a warning, a caution, etc., e.g. **u iat** (1) to a person in danger, you yourself, there is no one to look to but yourself, (2) You yourself, I shall look to no one but yourself, (3) Please yourself, do as you think proper, (4) You have no one to blame but yourself, you were the cause. **2.** Used as a comparative. **Ia ra ngala iat**, it was and is the larger. **3.** Is used with adverbs, very, still, early, as **ania** formerly, **ania iat** 'still further back; **a malana iat** 'the early morning'; **vailik iat** far, very far, **uro iat** [unglossed in the dictionary - GS]. **4.** It seems to refer to the past. **Damana iat**, it was like that before, originally (and is still); **utul a bung iat,** three days ago.

Lanyon-Orgill (1960) adds that in the case of *a malana iat*, the sense of *iat* is to be taken as 'real', the real or true morning being the early morning. Clearly, *iat* possesses the major meanings of the emphatic markers found in TP and the other New Guinea Austronesian languages we considered. As an intensifier, *iat* is used particularly with temporal and spatial expressions, and it is also a focus marker ('emphatic' or 'reflexive'). Lastly, the glosses given for the 'past reference' and 'comparative' uses would certainly encompass English

---

8   Mosel's interlinear glosses consistently label *iat* a particle, and in the introduction to her texts, she notes (1977:ix) that 'PART' means "emphatic particle". However, she rarely provides any gloss for these particles in her free English translations, and often I concur that forcing an English equivalent would be quite awkward. Examples (10) - (12) are the exception in this regard.

'still'. This provides the probable link for Tolai speakers' identification of their *iat* with English *yet*, which also means 'still', despite its more usual English collocation as a negative polarity item.

I examined a sample of Mosel's texts to see how Tolai use of *iat* parallels TP speakers' use of *yet*. In eight of these texts totalling 462 lines and representing five speakers, I found 54 instances of *iat*, with the various functions as shown in Table 3.

TABLE 3: INSTANCES OF TOLAI *iat* ACCORDING TO VARIOUS FUNCTIONS

Functions of *iat*	No. of cases
focus marker with noun:	23
focus marker with pronoun:	10
intensifier with locative/temporal:	16
other intensifier (adverbial):	2
other	3
Total	54

The major difference between the contemporary usage of Tolai *iat* and TP *yet* is that *yet* as a focus marker is specialised to pronouns, whereas *iat* occurs with both nouns and pronouns. Both, however, are used as intensifiers with locatives, temporals and other adverbials.

The puzzle of where *yet* came from and why it developed as it did seems to be solved. Speakers of Austronesian languages in the contact situation heard English *yet*, used mainly as a negative polarity item, but also sometimes as the positive 'still'. Tolai speakers identified it with their own *iat*, an intensifier also associated with the meaning 'still'. *Iat*, however, has focus-marking functions that English *yet* has not. *Yet* was pressed into service as a focus marker in Tok Pisin, and also, over time, drastically reduced its function as a negative polarity item. Its distribution in contemporary TP is split between adverbial intensification and the focusing of pronominals, much the way *iat* works in Tolai, and the way other emphatics, focusers, reflexives and limiters do in MAn languages, located to the right of the element they modify. It is not surprising that in the area of focus, with its great importance in rhetoric, the speakers of TP have shaped their language to fit the patterns their ancestors have used from *bipo yet*.

3.2 'OVERLAY' CONSTRUCTIONS IN NARRATIVE

Work on narrative texts in TP has revealed a discourse pattern that resembles the "overlay" type of construction discussed in the context of Papuan languages by Grimes (1972). See example (13). (Note that commas indicate phrasing of the speaker, and there is a clear 'full stop' intonation at the end of each line.)

(13) a. *Narapela man, em i-putim wanpela wetpela gras pisin.*   [Mr T.D.]
      other   man he put.on a   white   feather bird
      The one man put on a white feather.

    b. *Narapela em i-putim, blakfela, bilas.*
      other   he put.on black   dressed.up
      The other man put on a black feather, as a decoration.

    c. **Putim** *nau,*   *na* *i-go.*
         put.on PUNCT and go
         Having put them on, they/he went off.

Here, a verb is first used in an apparently finite context – the *i-putim* of (13a) and (13b) – but then carrying the story further, it is repeated in a context where it has a non-finite flavour, as in (13c), prior to the next finite verb that carries the narrative forward. I originally noted this pattern during my initial (1977) exploration of the cliticisation of *i-* (Sankoff [1977] 1980:265). My idea was that the context illustrated in (13c) would be a conservative one in terms of the generalisation of *i-*, which I thought was being grammaticalised as a finite verb marker. In further exploration of the patterning of *i-* (to be discussed in §5), it has turned out that this context does not significantly differ from other syntactic environments in terms of the presence of *i-*. It is, however, clearly distinctive in the absence of a subject pronoun in clause-initial position (as opposed, for example, to a position following coordination).

Looking into the history of this phenomenon, I found numerous examples in the speech of both Tagarak and Chavi, Sepik area speakers whose stories were transcribed by Gregory Bateson and Margaret Mead respectively, and published in Hall (1943). See examples (14) and (15). Note that in (15c), the non-finite form is actually overlaid on a noun, *skel*, rather than a previous verb.

(14) a. *Na Ainjang* **karim** *em* *i go daun.*   [Tagarak, in Hall 1943]
        and Ainjang carry him go   down

    b. **Karim** *i go na*   *putim ananit long haus* *bilong em.*
       carry   go and put   under LOC house POSS him
       And Ainjang carried him down. Carried him off and put him underneath his house.

(15) a. *Nau mi*   *stanap long mak, ol* *i* **makem** *mi.*   [Chavi, in Hall 1943]
        then 1SG stand LOC mark they mark   1SG

    b. **Makem** *mi* *finish, mi* *go sidaun long* **skel**.
       mark   1SG COMP 1SG go sit   LOC scale

    c. **Skelem** *mi* *finish, nau em* *i tok: "Orait, yu* *go daun, wok".*
       weigh   1SG COMP then 3SG say   alright 2SG go down work

       Then I stood at the mark and they measured me.
       After being measured, I went down to the scale.
       After weighing me, he said, "Alright, you go off to work".

These contrast with other examples that seem to have similar discourse structure but different syntax, such as (16), in which the repeated verb appears with both *i-* and subject pronoun:

(16) a. *Na* *Kowi i putim* *long nek* *bilong em,* *surukim* [Tagarak, in Hall 1943]
        and Kowi put   LOC neck POSS 3SG slip.tight

       em, *na* *i* **fasim**.
       3SG and fasten

    b. *Em* *i* **fasim** *nau,*   *na em* *i toktok long ol* *sisa* *bilong mitufela.*
       3SG fasten   PUNCT and 3SG talk   LOC PL sister POSS 1DU.EXC

And Kowi put it [the noose] on his neck, slipped it tight, and fastened it.
Having fastened it, he talked to our sisters.

The episode of Tagarak's story cited as (16) shares with the one from Mr T.D. (a Papuan language native speaker from Chimbu) in (13) the use of *nau* to punctuate or close off the action of the overlaid verb before continuing with the subsequent event. Keesing (1988:178) analyses Bislama *nao* as a perfect marker,[9] with "exactly the same force" as the perfect-marking particles in the Malaita languages, that is, they "articulate a state at a reference time (the time of the speech event) to an earlier state or event to indicate that the two are essentially and inseparably connected, and to focus attention on the present state". Following Simons (1985), Keesing notes that such a perfect marker is homonymous with a postposed topicaliser in many Malaitan languages, and that *nao* is also a topic marker in SIP. Most of the examples of *nao* in Keesing's texts do not co-occur with an 'overlaid' verb as in TP examples (13) - (16), serving instead to punctuate events each of which moves the action forward. In (17b), however, we see an example in SIP that is quite reminiscent of the TP overlaid structure (punctuation, interlinear glosses and English translation as in the original; I have begun new lines for each clause).

(17) a. *mifala*   ***go*** *faend-em* *nate,*   [Tome Kwalafane'ia, in Keesing 1988:245]
          SRP(weE)   go   find-TRS   nut

          ia?
          RHET

     b. ***googo*** *finis*   *nao*
          AUX   be.finished PRF

     c. *mifala*   *teke* *nate* *ia*   *kam* *nao,*
          SRP(weE)   take   nut   DEM   DEI   PRF

     d. *mifala*   *sidaon* *weit-em*   *olketa* *siton*   *nao...*
          SRP(weE)   sit   with-TRS   PL   stone   PRF

     We went to hunt for canarium nuts.
     And then we took those nuts.
     We sat down with stones...

This is an excerpt from Keesing's 1984 recording of Tome Kwalafane'ia, "a Kwaio man in his sixties who had worked on plantations in the late 1930s, then had served with the Solomon Islands Labour Corps in World War II" (p.231). In this excerpt, the instances of *nao* at the ends of the clauses in (17c) and (17d) serve to punctuate events in the story. In (17b), however (a line which is not glossed separately, but together with (17c) in Keesing's English translation), there is a partial repetition of the verb of (17a), which makes this *nao* seem functionally akin to the *nau* of the TP examples in (13c) and (16b), and *finish* in (15b) and (15c).

Another example from SIP, cited as (18), shows overlay quite parallel to the TP construction, with the second *kat-emu* occurring with no subject pronoun or SRP. Like the TP case in (14b), the speaker, Jonathan Fifi'i, uses the post-verbal *googo* auxiliary (see §5).

---

9   He also notes (following Simons 1985) that the same particle often serves as a topicaliser, following a noun or pronoun, a usage that is also common to SIP *nao*.

(18) a. *nao hem-i   kat-emu wan-fala*   [Jonathan Fifi'i, in Keesing 1988:233]
       so   SRP(he) cut-TRS one-ADJ

   fisi   long hemu.
   piece POSS   it

   b. *kat-emu googo finis,*
      cut-TRS AUX   be.finished

   c. *hem-i   safen-em   lelebeti bodi bulong hem.*
      SRP(he) shave-TRS   slightly body POSS   it

   So anyway, he cut one piece of it.
   He cut it and then
   he shaved down the casing slightly.

As far as Bislama is concerned, *nao* does not figure in the list of tense/mood/aspect (TMA) markers discussed by Crowley (1990a:211-212), though it clearly functions as a topicaliser as it does in SIP and TP. The existence of 'overlay' structures has not been discussed in the literature on either Bislama or SIP, to my knowledge. The several dozen pages of texts I have personally examined in SIP (in Keesing 1988) contain one or two other examples in addition to the two cited here; the Bislama texts I have looked at in the collection do not appear to contain close parallels to the TP examples which are fairly frequent in my texts.

My current tentative summary of the distribution of 'overlay' constructions in the Bislamic languages is that they appear to be more characteristic of TP than of the other two languages. Such constructions are not typical of English, and I would suggest that substrate influence is the likely source. I first thought that this influence might be from the Papuan languages, where medial verbs in narration do not appear fully inflected. Further checking, however, confirmed that this pattern predates the massive post-World War II spread of TP to native speakers of the Papuan languages of the New Guinea highlands. Attestations prior to World War II, when An influence can be considered to have been dominant, may point to MAn languages as models. The two Sepik area men from whose 1930s texts I cited examples were themselves native speakers of Iatmul, a Papuan language; however, they learned TP on the Gazelle Peninsula, almost certainly from speakers of MAn languages, in the 1920s. A more extensive examination of early texts as well as of likely substratum sources would be necessary to better understand the history of this development in TP. To what extent the influence of Papuan speakers (like Chavi, Tagarak and Mr T.D. in (14)), has been responsible for the spread of this discourse feature in TP is also not known. This is clearly a topic that merits further investigation.

4. INTERCLAUSE RELATIONS

In this section I attempt very briefly to set the record straight as to the analysis Penelope Brown and I (Sankoff & Brown 1976) proposed of relative clauses marked with *ia*, another area in which our original proposal posited a substrate origin for the beginnings of the grammaticalisation that we observed in TP. We suggested that the unstressed TP *ia* that is frequently postposed to nouns and pronouns has been, through frequency of usage, "downgraded" from a deictic to a determiner (a transition widely attested in the history of the world's languages). Occupying that slot, it was in a position to be pressed into use as a

bracketing device, for whatever kinds of constituents might typically follow a noun – other single nouns, phrases, or whole clauses, as exemplified in (19) - (21).

(19)   *Disfela ia,   ol   i-kosim em haumas?*   [Mrs Ta.]
          this   DEM 3PL charge 3SG how.much
          This one here, how much do they charge for it? (pointing to a piece of cloth)

(20)   *Em liklik barata ia   [mi tok ia].*   [Mrs M.S.]
          it   little brother DEM 1SG talk DEM
          It's the little brother I'm talking about.

(21)   *Dispela man ia,   [lek bilong en i-dai ia],   em*   [Mrs L.G.]
          this   man DEM leg of   his die DEM 3SG

          i-stap insait nau.
          stay   inside PUNCT
          This man whose leg was injured, he stayed inside.

In justifying our proposal that fluent second-language (pidgin) speakers innovated this strategy, we said that "this is particularly likely...because many Austronesian languages of the New Guinea and island Melanesian area show striking parallels" (Sankoff & Brown 1976:663). We illustrated the parallel first from Buang, and then cited Ray's 1926 work showing similar structures in Iai, Nguna, Tasiko, Uripiv and Tangoa.

Keesing's objection to this scenario is apparently twofold. Firstly, he says (1988:113) that the phenomenon is "not really a bracketing at all, since *ia* occurs [in all three daughter languages] in the same environments in the absence of the embedded clauses". This is quite puzzling to me, since this is the precisely the pattern we discussed at length in the original paper as the source of the later developments we observed in TP. Keesing goes on to dispute the idea that grammaticalisation of *ia* has involved stress reduction, stating (P.115) that "for Oceanic speakers the form has always been an unstressed *ia*". In N+*ia* constructions, whether or not involving embedded clauses, I would certainly agree that *ia* is likely always to have been unstressed. The stress reduction would have been between the putatively original 'place adverb' origin, with English source 'here' as originally discussed by Mihalic (1957), and the postnominal usage. In the 1976 paper, Brown and I said nothing about stress reduction in the N+*ia* environment, stating however that we did occasionally find instances of *ia* used as a place adverb. Nor did we suggest that the development from place adverb to postposed determiner occurred within TP itself. That N+*ia* constructions are found in nineteenth century sources in no way alters the scenario we painted. We cited evidence (Sankoff & Brown 1976:255) for the N+*ia* construction in both Bislama and SIP. That part of the development which seemed unattested in the other languages was the putting to use of the N+*ia* strategy in the service of relativisation.

Now, what we called *ia*-bracketing is a development we saw in 1971 in the speech of people from all over Papua New Guinea who were based in the Pacific Islands Regiment's Igam Barracks outside Lae. We doubted it was a 'Lae area' feature because of the regional diversity of our speakers and because of their personal histories of geographic mobility. Whether such a development has in fact occurred in TP's sister languages, or indeed whether it has survived and become more regular within Tok Pisin itself in the subsequent 20 years, I do not know. Whatever its fate, this innovation was clearly the contribution of second-language TP speakers whose mother tongue backgrounds predisposed them to find such a relativisation strategy congenial in another language. Brown and I used the development of

this relativisation strategy to argue that grammaticalisation is often spearheaded not by a first generation of native speakers, but by fluent second-language speakers in a speech community undergoing creolisation.

5. EVOLUTION OF AUXILIARIES IN TP

According to Crowley (1990a:201), "tense, aspect, and mood categories in modern Melanesian Pidgins represent one of those areas which is most noticeably different between the three varieties. These differences are therefore likely to represent innovations since the end of the recruiting era at the end of the nineteenth century". The onion-like system of auxiliaries I posited for TP (Sankoff 1984) differs from the other two languages particularly in its elaboration of the aspectual auxiliaries occurring to the right of the main verb. A revised view of the TP verb phrase is schematised as follows:[10]

$$\text{``T''- M} \qquad\qquad\qquad \text{A} \qquad\qquad\qquad\qquad \text{A} \qquad \text{``T''}$$

$$\text{(CLIT)} \ \text{(NEG)} \begin{Bmatrix} bin \\ ken \\ mas \\ nap \end{Bmatrix} \ save \ laik \begin{Bmatrix} go \\ kam \\ kamap \\ kirap \end{Bmatrix} \ V_1 \ (V_2) \begin{Bmatrix} igo \\ ikam \\ istap \end{Bmatrix} \dots pinis$$

The pre-verbal auxiliaries that derive mainly from English main verbs and modals are somewhat more similar to those occurring in the other languages. In Sankoff (1991) I describe some of the changes in the meanings of the irrealis marker bai, and also discuss semantic developments in the modals laik (volitional; also future) and ken (mainly deontic modal; specialised to the negative imperative). As yet, it is unclear whether any of these changes is related to substrate influence. Another modal, (i)nap, carries both deontic and epistemic meaning. It has apparently evolved into a 'raising' verb that can occur in a higher clause that embeds another clause.

As far as the syntax of bai is concerned, Suzanne Laberge and I suggested (Sankoff & Laberge 1973) that it had undergone the following changes: reduction from baimbai to bai; progressive loss of stress of the monosyllabic form; redundancy, in co-occurring with adverbs having future meaning; and word-order change (from sentence-initial to post-subject). Romaine (1990) has shown the dramatic continuation of the word-order changes through the mid 1980s. However, I am fully in agreement with Crowley (1990a:213-216) on Melanesian pidgin in general, and here disagree quite strongly with Keesing, that bai has not managed to become part of the verb phrase – try as it might. The observation originally made by Woolford (1977) still holds, and a sentence like (22) is ungrammatical.[11]

(22)   *Mi no bai kisim. [**bai** does not jump over neg; observation due to Woolford 1977]

A last pre-verbal element that appears to have undergone change that brings it into line with substrate patterns is the modal (i)nap (< Eng. 'enough'). Its use as a pre-verbal auxiliary, occupying the position that other modals such as laik and ken would also occur in, is illustrated in (23). It is also attested, however, as in (24), as a higher predicate under

---

10   This is a modified version of the schema originally presented in Sankoff (1984:113).

11   In discussing this point with Romaine in 1992, she said that although she had not specifically looked at negation across her own massive data base, she was fairly certain that she had not seen examples of bai occurring to the right of no.

which is embedded a complement clause, a pattern that has been observed in at least one substrate language.

(23)   *Ai, poroman, mi no nap i-kam arasait ia.*   [Mr D.]
       hey friend   I NEG able come outside DET
       Hey, buddy, I can't come outside!

(24)   *I-no nap   long   yutupela marit.*   [Mr D.]
       su loVBu be   melu   marit [cf.Buang]
       NEG able   COMP you.two marry
       It is not possible (i.e. permissible) for you two to marry.

TP apparently diverges most considerably from SIP and Bislama in the development of the post-verbal aspectual auxiliaries. Though TMA markers occurring to the left of the main verb are traditionally treated as auxiliaries, those occurring to the right have usually been treated as serial verbs. According to Goulden (1990:116):

> TP has a construction in which verbs of motion such as *go* 'go' and *kam* 'come' are used as directive co-verbs following the main verb, and comparable constructions are also found in Pijin and Bislama. *go* is used to denote movement away from the speaker and *kam* to denote motion towards the speaker...the locative verb *stap* can also be used in the *go/kam* slot, and indicates the place where something has come to rest, or within which a motion occurs.

He notes that serialisation of this sort is very common in the MAn languages, "many of which use a construction similar to that used by TP".

Looking at the data from Chavi and Tagarak, the view of *kam* and *go* put forward by Goulden seems to be supported by many examples:

(25) a. *Mi   wanfela tasol i stap nau,*   [Chavi, in Hall 1943]
        1SG alone   just   stay PUNCT

     b. *mi kalap i go long wata hia...*
        1SG jump go LOC water DET

        I was the only one left then,
        and I jumped into the water...

(26)   *Nau   mifela   ran stret   long pasich i kam.*
       PUNCT 1PL.EXC run straight LOC passage come
       So then we came running straight through the passage.

Moreover, "the place where something has come to rest" seems an apt description of (27), and of numerous other examples like it:

(27)   *Mi go sidaun long haus boi i stap.*
       1SG go sit   LOC house boy stay
       I went and stayed in the barracks.

Two other observations can be made about post-verbal *stap* in this period. Firstly, we note that in many cases, the original locative sense of 'stay' has been replaced by a more abstract sense of 'staying' (i.e. the notion of continuation). This is widely recognised in the literature on TP. Laycock (1970:xxii) states that *i stap* is a "common aspect marker...for continuous action". Wurm (1971:39) puts it this way:

Verbs followed by *i stap*...or preceded by *stap* indicate actions which are continuous, have started at a time which is before the point of time immediately preceding the one at which the action takes place, continue at the same level of intensity during the time referred to, and it is implied that the action will continue after that time, with indefiniteness concerning the length of time during which it will continue.

Mühlhäusler (1985a:378) identifies post-verbal *i stap* as a progressive. For example:

(28)   *Ol   i wok i stap   tudak.*   [Chavi, in Hall 1943]
        3SG work CONT night
        They kept working until night.

Secondly, *stap* is not only used in the post-verbal serial construction, but is also used frequently on its own, singly or iterated up to five times, to convey the idea that things continued pretty much the way they were, as illustrated in the underlined portions of (29) and (30).

(29)   *Em   i go stap wantaim nadafela big brada   bilong* [Chavi, in Hall 1943]
        3SG go stay with   other   big brother POSS

        em.   I stap.
        3SG stay
        He went and stayed with his other big brother. And there he stayed.

(30)   *Mi   stap long em. Stap stap stap stap stap*   [Chavi, in Hall 1943]
        1SG stay LOC 3SG stay stay stay stay stay
        I stayed on and on in it.

As far as *kam* and *go* are concerned, most observers note the 'directional' component mentioned by Goulden. According to Laycock (1970:xxiii):

Closely related to the...aspect markers are the directional markers *i kam* and *i go* (more common in Highlands Pidgin than in Lowlands Pidgin), which indicate whether a verb of motion describes an action which approaches the speaker or which goes away from him.

Both Laycock and Wurm, however, discuss another use of *i go*. According to Wurm (1971:45), *i go* is also used to indicate continuous action "even if no movement is implied". Wurm goes on to suggest a meaning difference between *i stap* and *i go* as continuative markers:

The difference between *i stap*...and *i go* is that the latter denotes actions which are expected to continue at equal intensity level for a considerable time after the time referred to in the clause, whereas in the case of *i stap*, indefiniteness concerning the duration of the action after that time is implied.

This subtle distinction does not seem to be borne out in the data I have examined. Rather, though *stap* is still used to some extent as a continuative, it appears to have been largely replaced by *go*, which now also occurs as an 'independent continuative' the way *stap* did in examples (29) and (30).The parallelism between the older use of *stap* and the modern use of *go* is illustrated by comparing example (31) from Chavi with (32) from Mona S., a woman from one of the three Labu villages south of Lae who was 22 when I recorded her in 1971.

(31) a. *Em   bilong Miwot,*   [Chavi, in Hall 1943]
       3SG POSS Miwot

     b. *nau   ol   imekim bigfela bilong em   long Kishit.*
        PUNCT 3SG make   big   POSS 3SG LOC Kishit

     c. *Nau   ol   ikatem skin bilong em   long Kishit.*
        PUNCT 3SG cut   skin POSS 3SG LOC Kishit

     d. *I stap. Nau   em   i bigfela,*
        CONT   PUNCT 3SG big

     e. *em   imarit, nau   ol   i tok...*
        3SG marry   PUNCT 3SG say

     He was from Miwot,
     but he was brought up in Kishit.
     They circumcised him in Kishit.
     And so it went. When he was grown,
     and he was married, they said...

(32) a. *Na ol   i go i stap long maunten ia*   [Mona S.]
        and 3SG go   stay LOC mountain DET

     b. *na ol   i-wet long bikpela guria   ia   i-go.*
        and 3SG wait LOC big   earthquake DET CONT

     c. *I go i go,   nogat.*
        CONT   NEG

     d. *Na ol   i kam bek   gen.*
        and 3SG come back again

     And they went and stayed on this hill
     and waited for the earthquake.
     Kept on [waiting], but it didn't happen.
     So they came back.

In (32b), Mona uses *go* to modify the stative verb *wet* 'wait'. She follows up in (32c) by using *i go i go* as an independent phrase, to indicate that this waiting continued for some time, in a way quite parallel to Chavi's use of *i stap* in (31d). Another example shows post-verbal *go* actually modifying the main verb *stap*:

(33)   *Mipela   stap igo   igo.*   [Mrs N.S.]
        1PL.EXC stay CONT   CONT
        We stayed on (and on and on).

What about *go, kam* and *stap* pre-verbally? Though both Wurm and Laycock mention that *stap* can occur pre-verbally with a meaning similar to its post-verbal 'continuative' meaning, the more than 230 instances of *stap* I examined in data from the 1930s, 1960s and early 1970s contain cases of its use as a main verb, a post-verbal auxiliary, and an independent continuative as in (31) above, but not a single case of pre-verbal *stap*. As far as *go* and *kam* are concerned, there has never been any suggestion of a 'continuative' meaning attaching to their use in the pre-verbal context. They seem simply to mean 'go and x' or 'come and x', as in (27) where Chavi 'went and stayed' in the barracks. In another example that shows both

pre-verbal and post-verbal *go*, the preverbal *go* can be glossed simply as 'went'. The post-verbal *go* indicates the progressive (i.e. the continuity of swimming as encoded in the English -*ing*). Far from meaning that they swam off in some direction, the sentence means they kept swimming around.

(34)   *Mipela   i go swim   i go.*   [Mrs N.S.]
        1PL.EXC go   swim   CONT
        We went swimming.

Occasionally *kirap* 'get up' (characterised by Mühlhäusler (1985a:178) as 'inchoative') or *kamap* 'arrive' will be used pre-verbally, meaning roughly 'up and do/did x'.

The view that post-verbal *go*, *kam* and *stap* in TP are parallel to, and possibly modelled on, the serial verbs in the MAn substrate languages is consonànt with understanding their primary meaning as being a directional adjunct to the main verb, with *stap* constituting a special 'stationary' case. Such systems are clearly found in relevant MAn languages, including Tolai. On the other hand, understanding post-verbal *go*, *kam* and *stap* as basically being continuous or progressive aspect markers, with *kam* and some uses of *go* also containing directional content, is a view conducive to their analysis as right-branching auxiliaries. I suggested in 1984 that a right-branching auxiliary system might be more consistent with the word order of SOV languages, and that perhaps the influx of Papuan substrate speakers in the past 50 years of TP history has helped promote the development of such a system. Wurm (1971) and Laycock (1970) both noted that this construction was more typical of the TP of the Highlands, where only Papuan languages are found.

If, however, *i*- is actually a 'predicate marker', as the classic view would have it, the fact that post-verbal *go*, *kam* and *stap* tend very strongly to occur with *i*- would mean that these are indeed 'predications' (i.e. a 'serial verb' view is the more reasonable one to adopt). And co-occur with *i*- they do, to such an extent that some authors (e.g. Wurm and Laycock) consistently refer to post-verbal *go*, *kam* and *stap* as *i go*, *i kam* and *i stap*, as we can see in the various quotations in the preceding discussion.

A quantitative study of the use of *i*- (Sankoff 1994) demonstrated that although *i*- is categorical in this context for the 'Premodern' speakers, this was not the case for the 1971 adult or child speakers I analysed. If the modern speakers have not retained *i* consistently in this context, does this mean they have reanalysed *i*, or does it mean they have reanalysed post-verbal *kam*, *go* and *stap*? This question will be taken up in the next section, in the context of a discussion of *i*.

Before proceeding to this last question, let us summarise the situation regarding the use of *kam*, *go* and *stap* after the main verb across the three Bislamic languages. Apparently there is some use of 'directional' *kam* and *go* in all three. For SIP, Keesing (1988:245) makes this point in connection with the post-verbal *kam* in example (17c) above: "For the common Oceanic directionals 'hither' and 'thither' (in Kwaio *mai* and *kau*), Pidgin uses *kam* and *go*. In Bislama and Tok Pisin, these directionals are rendered with serial clauses, *i kam* and *i go*". Crowley (1990b:70) argues that the directionals are one of two types of "genuine serialization" in Bislama, and he includes *stap* with these as a "posture type verb".

Bislama has shown one development from this construction type towards aspectual marking, in that *go* post-verbally is the source of the frozen form *gogo*. Crowley (1990a:212, 216) describes this as an iterative marker that has evolved since the 1950s. In SIP, Keesing (p.233) glosses post-verbal *googo* simply as 'AUX' (discussing the example

cited as (18b) above), and notes that "Fifi'i is using *googo finis* here...as equivalent to Kwaio *kee sui*". Though he glosses *sui* (variously as 'be finished' and 'then'), *kee* is also glossed simply as 'AUX', and it is unclear what its semantic contribution is: continuative perhaps more than iterative, from the context. In TP, the iterative marker is the pre-verbal *save*, but, as we have seen, both post-verbal *stap* and *go* have been described as markers of continuous aspect. Since, as most observers agree, they usually occur with *i*, there is the possibility that they have also been reanalysed as frozen forms in which initial *i* has no meaning or function (i.e. *igo*, *istap* rather than *i go* and *i stap*).

Whatever the best synchronic analysis of post-verbal *stap*, *go* and *kam* in TP, it is clear that, of the three Bislamic languages, it is in TP that this construction is most highly elaborated. Substrate influence from An languages is clear in the initial 'directional' stage that is shared by all three. What the unique circumstances are that have led TP to further elaboration in this direction are not entirely obvious, but may have to do with either the specifically Melanesian An languages with which it has been in contact longest, or with the later influence of Papuan languages.

6. MELANESIAN PIDGIN 'CORE GRAMMAR' AND THE QUESTION OF *i*-

At the heart of Keesing (1988) was a discussion of *i*-, referred to in much of the previous literature as a 'predicate marker'. Keesing analysed it as a "subject referencing particle" or SRP, built as a calque on the grammars of Oceanic languages. I had in fact independently proposed that *i*- could be regarded as a subject marker (Sankoff 1984), but suggested that pending a full-scale analysis of trends over time, and of the effects of creolisation, it was not clear how successive generations of speakers had in fact incorporated it into their grammars. In readdressing the question here, I will attempt to accomplish two goals.

Firstly, I will re-evaluate the picture I painted in my own earlier analysis (Sankoff 1977). I argued that *i*- had been cliticised from 'he' in a grammaticalisation process involving the reanalysis of a discourse pattern in which 'he'/*i*- was initially used in focusing subject nouns. Keesing provided a very fair-minded summary of my arguments, but took me to task for not seeing this construction as a calque on the Eastern Oceanic pattern of referencing nominal subjects with a pronominal copy in immediate pre-verbal position. Initially, then, I will use the data available to me for TP to evaluate Keesing's position on the origin of *i*- as a calque, and his analysis of it as a subject referencing particle rather than a predicate marker.

This debate, however, refers mainly to developments prior to the period for which our richest current data is available. Secondly, then, I will use materials from the twentieth century to discuss developments with respect to TP *i*-, which seems to have undergone considerable change over the past fifty years.

6.1 PREDICATE MARKER OR SRP: ORIGINS AND GRAMMATICAL STATUS

As with many academic controversies, most of the debate about *i*- has concerned matters on which the data are fairly thin: in this case, developments that occurred as part of the genesis and early evolution of Pidgin English in the Pacific in the nineteenth century. Scholars are reasonably agreed on the unity, prior to about 1880-85, of what came to be the three modern Bislamic languages, and agree that they began evolving separately at about this time. The ancestors of the present-day Bislama speakers, the Ni-Vanuatu of the then New

Hebrides, were the first to become involved in large numbers in the labour trade for plantation work, followed by Solomon Islanders and then Papua New Guineans. The early history of Beach-la-Mar is thus what must be investigated first, and on this, Crowley's extensive work is authoritative.

According to Crowley, the attested pronouns in early Beach-la-Mar (1840s-1860s) were as presented in Table 4. His thorough examination of the corpus available from that period leads him to observe (1990a:193) that, in the third person, "there was a consistent distinction between subject and object forms" and that, in the first person, "there is evidence for an optional separate subject form as distinct from the object form".

TABLE 4: ATTESTED PRONOUNS IN EARLY BEACH-LA-MAR (1840S-1860S)
[REPRODUCED FROM CROWLEY 1990A: HIS TABLE 5.1, P.193]

Person	Subject	Object
First Person	*mi, ae*	*mi*
Second Person	*yu*	*yu*
Third Person	*i*	*im*

Crowley (p.231) documents the introduction of *hem* as a subject in the 1870s-1890s period, during which time it was used less frequently than *i*. He continues (p.224):

> Between the turn of the century and the end of the 1920s, a number of additional
> pronominal categories are attested for the first time while the extent of variability
> in the system appears to have been reduced. The earlier variable use of *ae* and *mi*
> was resolved with the original object form completely replacing the subject form.
> In the third-person singular *hem* also replaced *i* as the canonical pronoun.

This is a view entirely consonant with my own (1977) reconstruction of developments through the 1920s. In addition, Crowley's research supports my analysis of the initial uses of the N+*i* construction as representing focus.

How does Keesing's picture differ from this one? Noting that Oceanic languages typically have two kinds of third person pronouns, focal pronouns and SRPs, Keesing argues that basically from the beginning, speakers of these languages associated the English 'he' [=*i*] with the Oceanic SRP, and tended to use it where their languages would dictate the use of SRPs. I had proposed that at the the time when 'he' [=*i*] was the subject pronoun, N+*i* structures were used only in a small minority of cases with full NPs, and seemed to have "contrastive force" (Sankoff 1980:262) or to be used with particularly long subject NPs. When *(h)em* replaced *i* as a subject pronoun, the N+*i* strategy became generalised as the normal unmarked pattern, losing its pragmatic force. Later we see the N+*em* sentences being used in similarly marked pragmatic circumstances. Keesing (1988:152) argues against this view, stating that:

> The far simpler analysis, simply on linguistic grounds, is to see *hem* as the
> pronoun that fits into the subject-NP slot, so that *hem i* sequences are
> grammatically parallel to, for example, *tana man i* in "Tanna man he lazy, he
> plenty lazy, he no like work" (Tanna, New Hebrides, 1867 – Adams 1984:174)
> - *even if hem i does not turn up in the early texts*. [my emphasis]

Perhaps this is a far simpler analysis, on linguistic grounds, but it is not quite in accord with the historical record, as Keesing himself admits. The first *(h)em+i* sequence I noted was in Churchill's 1911 compendium; the earliest citation Keesing locates is from Torres Straits

creole in 1898. Keesing cites turn of the century examples to illustrate that *mi mi* and *iu iu* sequences were being used, parallel to *em i*. Crowley, however, argues persuasively that these structures were always used in a minority of cases, and always represented focus, not the unmarked strategy.

Supporting my original view of the matter, then, is the sequence of attestations in the historical record, as well as the apparent pragmatic force of the examples in which we find pronoun doubling, inferrable not only from their semantic content but also from the fact that such examples always constitute a minority of cases. Let us assume for a moment, however, that it is simply the patchiness of the historical record that is responsible for the gap in early attestations of *(h)em* subject. Why would we not find N+*i* sentences occurring more frequently? According to Keesing (1988:145):

> If indeed what to English speakers was a "resumptive" and syntactically and semantically redundant "he" was being analyzed by Oceanic speakers as an SRP, what is crucial is in fact *not* the occurrence of this pronoun following noun subject, but the maintenance of subject reference pronominally in subsequent clauses. We have seen that in many Oceanic languages it is precisely in the clause where a noun subject is introduced that the otherwise obligatory SRP is optionally...deleted.

If the SRP is simply always optional with a subject noun, we have no explanation of why this construction was rare early on and became regular later. Moreover, there is an increased burden of argument on Keesing that there be a formal distinction between focal pronoun and SRP, such that when one finds in a text an item that could be either, it will be clear which it is. If focal pronouns and SRPs were all homonymous, postulating two separate series would indeed be unmotivated. Unfortunately, in most persons, such homonymy does exist. Keesing proposes the paradigm found in Table 5 for the pronouns of 1920s-1930s Solomons Pidgin:

TABLE 5: KEESING'S ANALYSIS OF SUBJECT PRONOUNS IN SIP OF THE1920S-1930S [HIS TABLE 12.2, 1988:192].

[Note: I have indicated those pronouns for which there is a formal difference between focal pronoun and SRP in bold.]

Number and person	FP	SRP
Singular 1	*mi*	*mi*
Singular 2	*iu*	*iu*
Singular 3	**hem**	**hem-i, i**
Dual 1 inclusive	*iumi(tufala)*	*iumi(tufala)*
Dual 1 exclusive	*mitufala*	*mitufala*
Dual 2	*iutufala*	*iutufala*
Dual 3	**tufala**	**tufala(-i), i**
Plural 1 inclusive	*iumi*	*iumi*
Plural 1 exclusive	*mifala*	*mifala*
Plural 2	*iufala*	*iufala*
Plural 3	**olgeta**	**olgeta(-i), i**

Only in the third person is there a formal distinction between the focal pronoun and the SRP. Thus, for example, if one finds a *mi mi* sequence in a text, one analyses the first as the focal

pronoun and the second as the SRP. If one finds only one such pronoun, it is, by definition, the SRP. Now what about the third person? Apparently for duals and plurals, if one finds *tufala-i*, or *olgeta-i*, this in itself represents only the SRP, and not the focal pronoun plus SRP. But the bracket around the (-*i*) in both these cases apparently means it is optional, and so either *tufala* or *olgeta* on its own *can* represent an SRP – thus erasing the formal distinction that would enable us to tell which was which other than by definition. To me, an additional problem here is that it is very unlikely that a form containing three or four syllables would cliticise and become part of the verb phrase (which is the way Keesing defines SRPs). Typologically, cliticisation tends overwhelmingly to involve unstressed monosyllables.

There remains only the third person singular case, where the SRP can apparently be either *i* or *hem-i*. Here, Keesing admits reanalysis of the surface strings he finds because in many cases *hem-i* appears in environments where his model predicts an SRP. Thus he must postulate that speakers have redefined *hem-i* as an SRP. At other times he simply finds *i* in the same environments, so it is also, as before, an SRP. Allowing these two options in apparent free variation certainly vitiates any attempt to explain the occurrence of *i* per se. However, we are provided with an environment in which *hem* would not be permitted (i.e. in immediate pre-verbal position). Thus one should not (and apparently does not) find cases of, for example, *hem go. Further, Keesing argues that in the 'nonverbal' sentences (locatives and equationals), where substrate languages show only focal pronoun subjects, and not SRPs (there being no verb for them to cliticise to), Solomons Pidgin shows the same pattern. The examples he adduces (p.168) are, however, nominal rather than pronominal, and so remain inconclusive. Although he cites (p.169) one example containing a pronoun from Pionnier's 1913 paper, this is also unclear since it is a dictionary definition not containing any predication. Nevertheless, it is on this point that I believe Keesing's model provides an interpretation of regularities observed by other scholars. It is well accepted that *i* in TP, for example, occurs less frequently with equationals than in other environments (e.g. Wurm 1971:13). It should be said, however, that a 'predicate marker' interpretation might equally well account for these facts: in the absence of a predicate, the predicate marker would also be absent.

Where does this leave us in interpreting *i*? To my mind, the bulk of the evidence still points to an interpretation according to which early early N+*i* constructions represent focus, just as later N+*em* constructions do, once *(h)em* had replaced *i* as subject pronoun. However, I do believe that the overwhelming evidence brought to bear by Keesing regarding the SRP pattern in Oceanic languages probably exerted a continuing influence on speakers to innovate a construction that leads to strings in the Bislamic languages resembling the output of a grammar containing SRPs. Indeed, this is likely to have been one of the forces leading to the generalizing of *i* to non-focused contexts, so that it eventually became the unmarked pattern.

In understanding the replacement of subject 'he' [=*i*] with *(h)em*, however, it is important to remember the social context. Melanesians in contact with Europeans in the 1860s, and to some extent the 1870s, had a much greater chance of learning English patterns because of the small scale of operations and the higher ratio of English speakers. Someone who worked on a small vessel fishing for bêche-de-mer, for example, would have a better opportunity to learn English than someone who was a member of a large labour line on a plantation. The greater salience of object pronouns of English would make them easier to learn on the basis of limited contact than the less salient subject pronouns. Thus *(h)em* and *mi* replaced *i* and *ae/ai* in subject position, thereby creating a pronominal system unmarked for case. Was 'he'

[=i] pushed into another status by its previous pronominal function being taken over by a different form? If i --> em was a 'push' chain because of the demographic factors in the pidginisation situation just outlined, it was also a 'pull' chain because of the SRP pattern Keesing demonstrates in the substrate languages. Indeed, it is likely that the substrate SRP pattern motivated many unstable experiments with pronominal systems during the nineteenth century, of the sort described by Keesing (1988:70-88). It is certainly possible that a construction analysed as FOCUS + SUBJECT by some speakers (as Crowley and I envision the early data) was analysed as SUBJECT+ SRP by others (as Keesing interprets these same data).[12]

In summary, I feel there is much to learn from Keesing's exposition of the pronominal systems of Eastern Oceanic languages, and would agree that they have been a continuing factor in the analysis Melanesians have made first of the English pronominal system, and then of the Pidgin spoken by other Melanesians from whom they learned it. If Keesing has not convinced me of the grammatical status of i as an SRP throughout the history of Bislama and of SIP, he has certainly brought into focus the potential the system carries for receiving differing analyses. It is to the reanalysis of i by generations of TP speakers since the 1920s that we turn in the next section.

6.2 THE REANALYSIS OF i IN TP (1920-70)

In a quantitative analysis of i in TP, the sample of speakers from whom I examined texts was chosen in order to look at change over time; differences among contemporary second-language speakers according to substrate (adult speakers of Austronesian versus Papuan languages); and creolisation. Table 6 shows that whereas the 'Premodern' speakers recorded by anthropologists in the early 1930s used i 65% of the time, this had decreased to 41% for the creole speakers of 1971.

TABLE 6: OVERALL USE OF i- BY FOUR GROUPS OF TP SPEAKERS

[Note: Only 2 or 3 cases of i- were found with the hundreds of mi, yu and yumi subjects examined, and so these forms were excluded from the tabulations. The proportion of use of i- was calculated for the following subjects: em, ol, all -pela forms, all full noun subjects, and zero-subjects. Post-verbal stap, kam and go did not vary by person of the subject, thus all instances were considered. 'Children' are the actual children of adults in the 'Coastal' and 'Highlands' samples.]

Group	No. of speakers	Substrate language family	Speaker generation	Approx. birth dates	Mean % i-	No. of tokens
'Premodern'	3	2 Papuan; 1 MAn	adult	1900-1915	65%	845
'Coastal'	4	MAn	adult	1931-1948	64%	777
'Highlands'	4	Papuan	adult	1935-1945	55%	1032
'Children'	6	(TP native)	child	1954-1964	41%	1289

It is clear from Table 7 that the post-verbal environment is the most favourable to the use of i-, for all speaker groups, though no group in the modern period has retained the

---

[12] Keesing's discussion of i- as a calque begins with his citing of Hall's view of the matter, according to which it "reflects a merger of the substandard English habit of recapitulating a subject by means of a pronoun...and the Melanesian-Micronesian feature of morphologically distinct pronouns that recapitulate subjects and introduce predicates, as in Marshallese ladrik e-gerabal 'the boy, he works'" (Hall 1966:83). Hall's original insight is still the key beginning point for an understanding of the evolution of i.

categorical system of the 'Premodern' speakers.[13] It is also clear, however, that the verbs *kam*, *go* and *stap* in *all* environments are highly likely to occur with *i-*. One possibility, then, is that *i-* with these verbs has simply been reanalysed as part of the lexical items themselves (i.e. that *kam*, *go* and *stap* are now represented in TP speakers' lexicons as *ikam*, *igo* and *istap*). Though younger speakers have reduced their usage of *i-* to 88% in this environment, they have reduced it even more drastically in all of the other environments, so that the grammatical environments appear to be more differentiated for them than for the older speakers.

TABLE 7: PERCENTAGES OF *i-* PRECEDING VARIOUS PREDICATE TYPES, FOR 8 ADULTS
AND 6 CHILDREN FROM THE 1971 SAMPLE

	Premodern		Coastal Adult		Highlands Adult		Children	
PREDICATE TYPE	%	N	%	N	%	N	%	N
post-verbal *go*, *kam*, *stap*	100%	(45)	97%	(64)	95%	(88)	88%	(115)
pre-verbal *go*, *kam*	93%	(29)	95%	(38)	94%	(32)	68%	(63)
main verb *go*, *kam*, *stap*	79%	(101)	90%	(168)	92%	(186)	72%	(265)
negation	76%	(45)	97%	(29)	97%	(35)	69%	(29)
other pre-verbal auxiliaries	92%	(61)	80%	(30)	75%	(48)	57%	(46)
main verb (**not** *g,k,s*)	52%	(482)	43%	(443)	33%	(640)	19%	(760)
TOTAL		(763)		(772)		(1029)		(1278)

A quantitative, multivariate analysis of the factors leading to retention of *i-* was used to test whether the most important factor in the retention of *i-* was the lexical item itself, or the position (pre-verbal versus main verb versus post-verbal). Results indicated that for the 'Premodern' generation there was a categorical difference between the post-verbal environment and the others; that for the 1971 adult generation, both Coastal and Highlands speakers, there was no significant difference according to position; and that the 1971 native speaker generation was again significantly more likely to use *i-* in the post-verbal environment than in the other two environments for *go*, *kam* and *stap*. Main verbs other than *kam*, *go* or *stap* were already the least likely category for the use of *i-* among the Premodern speakers, and their usage shows a progressive decline of about ten percentage points across each of the four categories. Fewer than one in five main verbs other than *kam*, *go* or *stap* was used with *i-* by the child speakers. More detailed results from the analysis of ten different factors, including person of the subject, syllable structure of the predicate, and position in the string are discussed in Sankoff (1994), however these details will be recapitulated here only in so far as they help us to understand (a) the analysis of post-verbal *kam*, *go* and *stap*, an unresolved issue from the previous section of this paper; and (b) the issue of the analysis of *i-* as 'predicate marker' versus SRP.

From the materials we have at hand, it looks as if the analysis of *i-* has been a problem for TP speakers, as well as for linguists, during this period. Before proceeding to outline my interpretation of the analyses of *i-* made by successive generations of speakers, it is important to look at one more set of data, showing the overall frequency of *kam*, *go* and *stap* as a proportion of all verbs used by the speakers. Table 8 shows that whereas *go* constituted only

---

13   When I began this study, I had thought this categoriality might be an artefact of the conditions under which the texts were recorded from Chavi and Tagarak – the relatively slow speech of someone whose words are being transcribed in longhand. However, this was also true of the third member of this group: the older speaker from New Britain whose tape recording I transcribed myself.

roughly one in ten verbs used by the Premodern speakers, contemporary adults used it more frequently. For the child speakers, more than one in five of all the verbs used was *(i)go!* Their use of *go* particularly increased in the 'post-verbal auxiliary' category. Whereas for contemporary adults, the post-verbal category consists of a ratio of 2:1:1 for *go*, *stap* and *kam* respectively, the ratio for the children is 4:1:1.

TABLE 8: *Go*, *kam* AND *stap* AS A PERCENTAGE OF ALL VERB TOKENS IN THE CORPUS

	PREMODERN	COASTAL & HIGHLANDS	CHILDREN
go	11%	15.0%	22%
kam	7%	10.5%	7%
stap	5%	6.5%	6%
Total	23%	32 %	35%
N = ALL VERBS	(801)	(1809)	(1289)

My reading of the data is as follows: by the 1920s-1930s, the period I am here calling 'Premodern', *i-* had already been reanalysed as a predicate marker, according to the scenario outlined in §6.1. For the 'Premodern' speakers, post-verbal *i-go*, *i-stap* and *i-kam* were indeed predications (i.e. serial constructions), no doubt modelled on the widespread MAn substrate pattern. For them, *i-* was more clearly associated with the third person than it is now (Sankoff 1994). They had also begun, in part probably for phonological reasons, to show some weakening of *i-*, such that their pattern even in the third person was variable, rather than categorical. The generation that followed them began to increase its overall use of *kam*, *stap* and particularly *go*, as shown in Table 8. As *go* took on more aspectual than directional meanings, it became more closely associated with the main verb and was reanalysed as a continuative marker. What was for the 'Premodern' generation a serial verb construction was not seen as a serial construction any more, so something other than 'predication' was made of the *i-* that was so obviously associated with these verbs in the speech of their elders. This generation, represented here by both Coastal and Highlands adults of the early 1970s, apparently associated the *i-* with the verbs themselves, and it looked as if *go*, *kam* and *stap* were simply becoming lexicalised as *igo*, *ikam* and *istap*. Their children, however, in carrying the overall deletion process further, have retained the post-verbal environment as the main stronghold of *i-*, making their speech look once again like that of their ancestors some fifty years earlier, as if it were the output of a grammar in which post-verbal *i-go*, *i-stap* and *i-kam* were serial verbs.

Only an analysis that is driven by a theoretical bias towards the most abstract and general functions of grammatical particles would assign the role of 'predicate marker' to *i-* for the creole speakers. The quantitative analysis of current trends shows that this function, if it existed, is being lost or has been lost already.

7. THE SUBSTRATE REVIEWED

What can we conclude about the influence of substrate languages on TP? Firstly, there can be no doubt that three features of An grammar are reflected in TPs current development: the grammaticality of inalienability; the use of postposed focus particles; and the innovation in

relative clauses to use the postposed determiner as an opportunity for bracketing. In the TMA system, we see similar reflections of an An influence in the developments with respect to modals, in the shift of *bai* to iterative and punctual aspect marking, and in the evolution of post-verbal auxiliaries, which may reflect the influence of Papuan speakers.

Given this powerful evidence for substratal effects, it is surely appropriate that this paper be dedicated to the memory of Roger Keesing. His intimate knowledge of the Austronesian languages of the Solomon Islands, coupled with his keen observations on the process of language contact, led him to a fierce advocacy of substratal influence across the board. Although his approach to data was firmly rooted in the traditional paradigm of analysis by selective examples, and his theoretical drive led him to neglect the role of internal evolution in the Bislamic languages, his basic insight emerges almost unscathed.

Finally, we must be puzzled by the erosion of the *i*- marker with the new generation. Here we must directly confront the mystery which for me motivated this work, an instance of a development which must challenge all linguists: the fact that grammatical particles disappear as soon as they are created. This phenomenon appears in the history of many languages, but the rapid evolution of TP brings it to our attention over and over again. The preposition *long* is reduced to *l*, the auxiliary *save* to *sa*, *laik* to *la* or syllabic *l*, *baibai* to *b*. These reductions are only the obvious, audible evidence of the attrition process. The further reduction to zero, direct or indirect, is a normal consequence, and the subsequent morphologisation (or grammaticalisation) of these zeros leads to a reanalysis of the original system. It is easy to account for this process by a facile evocation of principles of least effort and the desire to speak quickly. But our whole understanding of the process of grammaticalisation as the creation of grammar is challenged by the finding that grammar can disappear as soon as it is created. The resolution of this mystery will most likely be found only through detailed studies of grammar in use through large-scale observations of face to face interaction.

REFERENCES

Adams, R., 1984, *In the land of strangers: a century of European contact with Tanna, 1774-1874.* Pacific Research Monographs 9. Canberra: Australian National University.

Camden, William G., 1979, Parallels in structure of lexicon and syntax between New Hebrides Bislama and the South Santo language as spoken at Tangoa. *Papers in pidgin and creole linguistics* No.2, 51-117. *PL*, A-57.

Capell, Arthur, 1969, *A survey of New Guinea languages.* Sydney: Sydney University Press.

Churchill, W., 1911, *Beach-la-Mar: the jargon or trade speech of the western Pacific.* Washington, D.C.: Carnegie Institution.

Clark, Ross, 1979-80, In search of Beach-la-Mar: towards a history of Pacific Pidgin English. *Te Reo* 22-23:3-64.

Crowley, Terry, 1990a, *Beach-la-Mar to Bislama: the emergence of a national language in Vanuatu.* Oxford: Clarendon Press.

1990b, Serial verbs and prepositions in Bislama. In Verhaar, ed. 1990:57-89.

Faraclas, Nicholas, 1988, Rumours of the demise of Descartes are premature. *Journal of Pidgin and Creole Languages* 3:119-135.

Goulden, Rick J., 1990, *The Melanesian content in Tok Pisin. PL*, B-104.

Grimes, Joseph E., 1972, Outlines and overlays. *Language* 48:513-524.

Hall, Robert A., Jr., 1943, *Melanesian Pidgin English: grammar, texts, vocabulary.* Baltimore: Linguistic Society of America.

1966, *Pidgin and creole languages.* Ithaca, N.Y.: Cornell University Press.

Jourdan, Christine, 1985a, Creolisation, nativisation or substrate influences: what is happening to BAE in Solomon Islands Pijin. *Papers in pidgin and creole linguistics* No.4, 67-96. *PL*, A-72.

1985b, Sapos Iumi Mitim Iumi: urbanization and creolization in the Solomon Islands. PhD thesis, Australian National University.

Keesing, Roger M., 1985, *Kwaio grammar*. *PL*, B-88.

1988, *Melanesian Pidgin and the Oceanic substrate*. Stanford: Stanford University Press.

Lanyon-Orgill, Peter A., 1960, *A dictionary of the Raluana language (New Britain, S.W. Pacific)*. Victoria, B.C.: The author.

Laycock, Don, 1970, *Materials in New Guinea Pidgin (Coastal and Lowlands)*. *PL*, D-5.

Lichtenberk, Frantisek, 1980, A grammar of Manam. PhD dissertation, University of Hawaii.

Lynch, John, 1973, Verbal aspects of possession in Melanesian languages. *Oceanic Linguistics* 12:69-102.

Methodist Overseas Missions, 1964, *A Kuanua dictionary*. Rabaul: Methodist Mission Press.

Mihalic, Francis, 1957, *Grammar and dictionary of Neo-Melanesian*. Westmead, NSW: The Mission Press.

Mosel, Ulrike, 1977, Tolai texts. *Kivung* 10/1-2.

1980, *Tolai and Tok Pisin: the influence of the substratum on the development of New Guinea Pidgin*. *PL*, B-73.

Mosel, Ulrike and Peter Mühlhäusler, 1982, New evidence of a Samoan origin of New Guinea Tok Pisin (New Guinea Pidgin English). *Journal of Pacific History* 17:166-175.

Mühlhäusler, Peter, 1976, Samoan Plantation Pidgin English and the origins of New Guinea Pidgin: an introduction. *Journal of Pacific History* 11:122-125.

1978, Samoan plantation Pidgin English and the origin of New Guinea Pidgin. *Papers in pidgin and creole linguistics* No.1, 67-119. *PL*, A-54.

1985a, Internal development of Tok Pisin. In S.A. Wurm and P. Mühlhäusler, eds *Handbook of Tok Pisin (New Guinea Pidgin)*, 75-166. *PL*, C-70.

1985b, The number of Pidgin Englishes in the Pacific. *Papers in pidgin and creole linguistics* No.4, 25-51. *PL*, A-72.

1987, Tracing predicate markers in Pacific Pidgin English. *English World-Wide* 8:97-121.

1990, On the origins of the predicate marker in Tok Pisin. In Verhaar, ed. 1990:235-249.

Oxford English dictionary, 1994, On-line version, accessible through the University of Pennsylvania Library System.

Pionnier, J.-N., 1913, Pigeon English ou Bichelamar. *Revue de Linguistique et Philologie comparée* 46:109-117, 184-198.

Ray, Sidney H., 1926, *A comparative study of the Melanesian island languages*. Cambridge: Cambridge University Press.

Romaine, Suzanne, 1990, Change and variation in the use of *bai* in young children's creolized Tok Pisin in Morobe Province. In Verhaar, ed. 1990:187-203.

1993, The decline of predicate marking in Tok Pisin. In F. Byrne and J. Holm, eds *Atlantic meets pacific: a global view of pidginization and creolization*, 251-260. Amsterdam: Benjamins.

Sankoff, Gillian, 1977, Variability and explanation in language and culture: cliticization in New Guinea Tok Pisin. In M. Saville-Troike, ed. *Linguistics and anthropology*, 59-73. Washington, D.C.: Georgetown University Press.

1980, *The social life of language*. Philadelphia: University of Pennsylvania Press.

1984, Substrate and universals in the Tok Pisin verb phrase. In D. Schiffrin, ed. *Meaning, form, and use in context: linguistic applications*, 104-119. Washington, D.C.: Georgetown University Press.

1985, Touching pen, marking paper: Queensland labour contracts in the 1880s. In Deborah Gewertz and Edward Schieffelin, eds *History and ethnohistory in Papua New Guinea*, 100-126. Oceania Monograph No. 28. Sydney: University of Sydney.

1990, The grammaticalization of tense and aspect in Tok Pisin and Sranan. *LVC* 2:295-312.

1991, Using the future to explain the past. In F. Byrne and T. Huebner, eds *Development and structures of creole languages: essays in honor of Derek Bickerton*, 61-74. Amsterdam: Benjamins.

1993, Focus in Tok Pisin. In F. Byrne and D. Winford, eds *Focus and grammatical relations in creole languages*, 117-140. Amsterdam: Benjamins.

1994, An historical and evolutionary approach to variation in the Tok Pisin verb phrase. *Parasession on Variation and Linguistic Theory*. Chicago Linguistic Society 30.

Sankoff, Gillian and Penelope Brown, 1976, The origins of syntax in discourse: a case study of Tok Pisin relatives. *Language* 52:631-666.

Sankoff, Gillian and Suzanne Laberge, 1973, On the acquisition of native speakers by a language. *Kivung* 6:32-47.

Sankoff, Gillian and Claudia Mazzie, 1991, Determining noun phrases in Tok Pisin. *Journal of Pidgin and Creole Languages* 6:1-24.

Senft, Gunter, 1986, *Kilivila: the language of the Trobriand Islanders*. Berlin: Mouton de Gruyter.

Simons, Linda, 1985, Malaitan influence on two grammatical particles in Solomon Islands Pijin. *Papers in pidgin and creole linguistics* No.4, 53-65. *PL*, A-72.

Tryon, Darrell, 1987, *Bislama: an introduction to the national language of Vanuatu. PL*, D-72.

Verhaar, John W.M., ed., 1990, *Melanesian Pidgin and Tok Pisin*. Amsterdam: Benjamins.

  1991, The function of *i* in Tok Pisin. *Journal of Pidgin and Creole Languages* 6:231-267.

Woolford, E., 1977, Aspects of Tok Pisin grammar. PhD dissertation, Duke University.

Wurm, S.A., 1971, *New Guinea Highlands Pidgin: course materials. PL*, D-3.