Pieter Muysken

Beryl Bailey writes in her conclusion to *Jamaican Creole Syntax* (1966: 144): "I have demonstrated ..., that the language in question is a genuine Creole language." In saying this she does not develop any specific criteria for Creole languages, but argues that Jamaican is not just a "a dialect of English". Why is a given language classified as a Creole? Are only historical criteria used, so that a Creole would be something like a language spoken by people whose ancestors were slaves? Are there any specific and conclusive linguistic criteria for classifying a language as a Creole? This paper will try to contribute to an answer to this very general question, discussing it in relation to Papiamentu.

Bailey enumerates a number of differences between Jamaican Creole (JC) and English:

(a) no subject-verb agreement in JC;
(b) JC tense restricted to a past marker;
(c) JC has no passive;
(d) JC has a separate locative 'be' and an equational 'be';
(e) JC predicate adjective constructions have no copula;
(f) JC has both an aggregate and an associative plural; in English this is limited to first and second persons;
(g) JC generic phrases have no article;
(h) JC has no case distinctions in nouns and pronouns, nor a gender distinction in third person pronouns;
(i) The inverted sentence type is basic to Creole in order to indicate emphasis.

Curiously enough, most of these differences are negatively phrased: European Model X has what Creole X' has not. This perspective was typical of early descriptions of Creole languages. Often the alleged "lack of distinctions" is felt as a negative characteristic, only rarely as something positive, as in Lenz' book on Papiamentu (1928). Lenz takes Jespersen's point of view (Jespersen, 1928: 206) that "any simplification, any discarding of old superficial distinctions is progressive, ..." Viewing the disappearance of inflection in Creoles as a simplification presupposes the more or less conscious adoption of the structuralist framework in grammar: the classification of items in their various shapes. In a generative framework, the loss or lack of inflection often leads to greater complication of syntactic constructions (cf. Bever & Langendoen, 1972). A number of transformations which are crucially dependent upon the previous (cyclical) application of agreement and concord transformations cannot apply.

This paper will investigate various properties of relative clause formation, question formation and topicalization in Papiamentu, and present these properties within the framework of a general theory of creolization. In particular, the view will be brought forward that in the process of Creole formation, certain superficial similarities between different ancestor languages (European, African) have led to the restructuring characteristic of the Creole languages themselves.
Since "Wh-movement" plays a central part in the argument, it will merit separate discussion, preceded by some introductory remarks about the processes involved in creolization, and a brief description of the interaction of particular source and lexifier languages in the formation of Sénégal Creole. The main part of the paper will be devoted to various types of movement transformations, occurring in Papiamentu x/y questions, restricted relatives, and topicalization. Finally, the question raised in this introduction will be taken up again, with regard to the movement rules described.

I will take the term "creolization" here in its widest sense, i.e. as referring to the complex of processes of language contact, pidginization, acquisition by native speakers, and grammar expansion. I will have nothing to say about the issue of monogenesis and polygenesis, and implicitly reject the 'baby talk hypothesis' in its various manifestations. Instead, I will assume that both general processes of second language learning, processes of interference, and processes of autonomous grammar construction play a part in creolization. The resulting grammars of creole languages are then assumed to be the product of a compromise between specific African and specific European language structures, mediated through the psychological processes of second language learning and the specific linguistic capacities of grammar construction.

The interaction between the source and lexifier languages: the case of Sénégal Creole

An example of the complexity of interaction between the European lexifier languages and the African source languages will be given here from the pidgin (or Creole) spoken in parts of Sénégal (Chataigner, 1963). It involves three verbal suffixes: /-du/ 'passive', /-ba/ 'perfective aspect' (?), and /-ta/ 'causative'.

Passive /-du/ we find in sentences such as:

(1) kéla sabi ntchè-dù
that know understand-PAS
'that can be understood'

(2) así na múdu i tê gîtis ki ka-sina-dù
thus in world there are people that NEG-teach-PAS
'thus there are people in the world who are not taught'

While /-du/ formally derives from the Portuguese past participle marker -do, as in vido 'seen' or tido 'had', its syntactic use in Sénégal Creole reflects the use of similar passive markers in African languages which can be reasonably thought to have provided a substratum for Sénégal Creole, such as Fula:

(3) 'o-najjeiì gujjo 'on
'he caught the chief'

(4) gujjo 'on nanggama
'the chief has been caught' (Arnott, 1970: 240)

The past tense or perfective aspect marker /-ba/ can be traced directly to Portuguese -va, as in estava 'was' or pesava 'weighed'. Its precise meaning in Sénégal Creole is not clear, however, and it may well correspond to a segment of the exceedingly complex tense system of Fula and of related languages (cf. Arnott, 1970).
The causative marker /-ta/ we find in Sénégal Creole pairs such as qora 'cry' and qora-ta 'make cry'. Compare (5) and (6):

(5) bibus na gora
   alive ASP cry
   'the ones alive are crying'

(6) kë ki ka-ta-qora-ta si fiju, amanhà si fiju ta-qorn-ta-
   who that NEG-ASP-cry-CAUS his son, tomorrow his son ASP-cry-CAUS-PH
   'who does not make his son cry, tomorrow his son will make him cry'

Causative markers of this sort do not exist in Portuguese, nor does /-ta/ carry any resemblance to a Portuguese form (except remotely to estar 'be'). We do find causatives in several of the languages of Sénégal and neighbouring countries, e.g. in Mandinka:

(7) faa 'be full'
    fandi 'fill'

(8) bàtáa 'be tired'
    bátandi 'tire, worry'  (Rowlands, 1969: 104)

Also in Fula:

(9) 'ekkito 'learn'
    'ekkintona 'teach'

(10) seyo 'rejoice'
     seyna 'please'  (Arnot, 1970: 34)

While the causativizing suffix /-ta/ is not directly derived from these languages, the claim could be made that the general process of causativization in Sénégal Creole resembles that of other languages in the area. An additional example of the suffix is provided in (11):

(11) fiju malkriadu pode brynà-ta rásà
    son badly educ. can shame-CAUS kin
    'a badly educated son can put his kin to shame'

Here brynà 'be ashamed' is causativized through addition of the suffix /-ta/.

Summarizing, we may state that the three verbal suffixes stand in a very complex and pluriform relationship to both source and lexifier languages. It can be schematically presented as follows:

<table>
<thead>
<tr>
<th>FORM</th>
<th>MEANING</th>
</tr>
</thead>
<tbody>
<tr>
<td>/-du/</td>
<td>Portuguese</td>
</tr>
<tr>
<td>/-ba/</td>
<td>Portuguese</td>
</tr>
<tr>
<td>/-ta/</td>
<td>?</td>
</tr>
</tbody>
</table>

In the lexicon and in morphology, the interaction between source and lexifier languages is often conceived as a type of fusion between 'European' forms and 'African' meanings. In syntax, it is often harder to discern what is form and what is meaning, as will be seen in the analysis of Papiamentu.

"Wh-movement"

The languages which have contributed a considerable part of the lexicon of the Creole languages, such as Portuguese, Spanish, English, French, and Dutch,
evidence a process which has been described for English as Wh-movement (cf. Chomsky, 1976). Elements which carry the feature \(+Wh\) are moved to an initial position in the cyclic node, least controversially, in \(S\):

\[
\begin{array}{c}
\text{S} \\
\downarrow \\
\text{COMP} \\
\leftarrow \text{[+Wh]}
\end{array}
\]

Thus we find proposed derivations such as:

\[
\text{COMP} \rightarrow \left[ \begin{array}{c} \text{S} \\
\text{you are how} \end{array} \right] \rightarrow \left[ \begin{array}{c} \text{S} \text{ how [ } \text{S} \text{ are you ---?] \end{array} \right]
\]

Similarly in relative clauses we find:

\[
\text{the man} \rightarrow \left[ \begin{array}{c} \text{S} \\
\text{I saw [+Wh]} \text{ in the street} \end{array} \right] \rightarrow \left[ \begin{array}{c} \text{S} \text{ who [ } \text{S} \text{ I saw --- in the street] \end{array} \right]
\]

There is considerable disagreement regarding the question of where \([+Wh]\) actually moves to: either it moves into COMP, or to a position immediately to the left of COMP, or it is adjoined to COMP immediately to its left. This problem need not concern us here, and will not be raised in this paper.

In some languages, such as English, the \([+Wh]\) element can move across several clause boundaries, so that sentences such as (15) are possible:

\[
\text{The man who I think that Mary said that John saw --- in the street...}
\]

Here who refers to the underlying direct object of saw, several clauses downwards. Again, in these languages there is considerable disagreement whether the \([+Wh]\) element "hops from COMP to COMP" or whether it moves across variables freely. Here the framework proposed in Chomsky (1976) will be adopted for reasons of expository clarity.

An equally complex issue, more central to the arguments presented here, concerns the domain of "Wh-movement": in which constructions does it play a part? Here, too, I will follow Chomsky's line of argument: constructions which have the same general characteristics should be accounted for in the same way. "Wh-movement" in English is supposed to have four characteristics:

(a) it leaves a gap;
(b) where there is a bridge, i.e. where it applies across S-boundaries, there is an apparent violation of subjacency, the Propositional Island Constraint, and the Specified Subject Constraint;
(c) it observes the Complex Noun Phrase Constraint;
(d) it observes Wh-islands constraints.

I will analyze the grammatical constructions mentioned above: x/y questions, relative clauses and topicalization or clefting, to see whether they involve Wh-movement or not, using criteria (a)-(d).
Papiamentu x/y questions

In Papiamentu embedded and non-embedded x/y questions are formed in the same way: by fronting the questioned element. We find the following +Wh forms:

\[
\begin{align*}
(16) & \quad \text{ken} & \text{who} \\
& \quad \text{kiko} & \text{what} \\
& \quad \text{kwa} & \text{which} \\
& \quad \text{unda} & \text{where} \\
& \quad \text{kon} & \text{how} \\
\end{align*}
\]

Thus we find simple questions such as:

\[
\begin{align*}
(17) \quad \text{ken a bai fiesta} & \quad \text{who ASP go party} \\
& \quad \text{who has gone to the party?'} \\
(18) \quad \text{kiko bo ta tin den man} & \quad \text{what you ASP have in hand} \\
& \quad \text{what do you have in your hand?'} \\
(19) \quad \text{kwa homber bo a duna e buki} & \quad \text{which man you ASP give the book} \\
& \quad \text{which man did you give the book to?'}
\end{align*}
\]

And embedded questions:

\[
\begin{align*}
(20) \quad \text{Maria no a bisa ken a bai kas} & \quad \text{Mary not ASP say who ASP go home} \\
& \quad \text{Mary has not said who has gone home'} \\
(21) \quad \text{mi no ta sabi kiko bo ta tin den man} & \quad \text{I not ASP know what you ASP have in hand} \\
& \quad \text{I don't know what you have in your hand'}
\end{align*}
\]

The questioned element is obligatorily questioned when it is an NP', and optionally in the case of PP's. Both the set (a) and (b) are grammatical in the following examples:

\[
\begin{align*}
(22a) \quad & \text{kiko bo ta hunga kuné} & \quad \text{what you ASP play with-PRO} \\
& \quad \text{what are you playing with?'} \\
& \quad \text{b) ku kiko bo ta hunga} & \quad \text{with what you ASP play} \\
(23a) \quad & \text{ken b'a reken riba djé} & \quad \text{who you-ASP count on-PRO} \\
& \quad \text{who had you counted on?'} \\
& \quad \text{b) riba ken b'a reken} & \quad \text{on who you-ASP count} \\
(24a) \quad & \text{ken bo ta kumpra flor pè} & \quad \text{who you ASP buy flower for-PRO} \\
& \quad \text{who are you buying flowers for?'} \\
& \quad \text{b) pa ken bo ta kumpra flor} & \quad \text{for who you ASP buy flowers}
\end{align*}
\]
The fronting of Wh words can occur across S-boundaries, constituting an apparent violation of Subjacency, the Propositional Island Constraint, and the Specified Subject Constraint, in the same way as Wh movement in English:

(25a) **unda e boto ta pasa banda dję**
where the boat ASP pass along-PRO
'what route does the boat take?'

b) **banda di unda e boto ta pasa**
along of where the boat ASP pass

(26) **kwa seru b'a bisa ku b'a drumi riba dję**
which hill you-ASP say that you-ASP sleep on PRO
'on which hill did you say that you have slept?'

(27) **ken bo ta kere ku ta parse mi tata**
who you ASP believe that ASP look like my father
'who do you believe looks like my father?'

(28) **unda bo ta kere ku Wan a bisa ku Maria tabata biba**
where you ASP believe that John ASP say that Mary ASP live
'where do you believe that John said that Mary lived?'

(29) **ken Wan a ganjabo ku e a laga drenta**
who John ASP lie-you that he ASP let enter
'who did John lie to you that he had let in?'

It does observe the Complex Noun Phrase Constraint:

(30) **șken Wan a sakabo mentira ku e a laga drenta**
who John ASP tell-you lie that he ASP let enter
'who has John told you a lie that he has let in?'

(31a) **șkiko b'a faz e kwentu ku b'a hunga kuné**
what you-ASP make up the story that you-ASP play with-PRO
'what did you make up the story that you played with?'

b) **șku kiko b'a faz e kwentu ku b'a hunga**
with what you-ASP make up the story that you-ASP play

And also the Wh-island constraints:

(32a) **șken Wan a puntra su mes si e a lagé drenta**
who John ASP ask his self if he ASP let-PRO enter
'who did John wonder whether he let come in?'

b) **șera la laga**
let-Ø

Thus the movement which occurs in Papiamentu x/y questions conforms in all respects to characteristics (b), (c), and (d), given by Chomsky for Wh movement, and in some respects to (a). The problem of the /e/ pronoun which occurs with stranded prepositions will be discussed later.

**Relative Clause Formation**

Relative clauses in Papiamentu are mostly introduced by an invariant marker /ku/. When the relativized NP is subject, direct object, or indirect object of the embedded clause, no pronominal element is present:
When the relativized NP is part of a prepositional phrase, the preposition and a pronominal element remain stranded in the PP position:

(36) e kas ku mi ta biba banda djé ...
    the house that I ASP live next to-PRO
    'the house that I live next to'

(37) e pen ku m'a skirbi kuné ...
    the pen that I-ASP write with-PRO
    'the pen that I wrote with'

(38) e sen ku mi ta warda riba djé ...
    the money that I ASP wait on PRO
    'the money that I am waiting for'

Note that the PRO element cannot be deleted from the PP here:

(39a) e homber ku mi amigu ta bai merka kuné ta di Korsow
    the man that my friend ASP go market with-PRO be from Curacao
    'the man that my friend goes to the marked with is from Curacao'

b) *ku-∅

Examples (41) and (42), involving a plural relativized noun, demonstrate the invariability of the PRO element, in contrast to "normal" pronouns, which are distinguished according to number, as in (40):

(40a) e ta di Korsow
    he be from Curacao
    'he is from Curacao'

b) nan ta di Korsow
    they be from Curacao
    'they are from Curacao'

(41) e hombernan ku mi amigu ta bai merka{ku-nan kuné} ta di Korsow
    the man-PL that my friend ASP go market{with-them} be from Curacao

(42) e serunan ku mi amigu ta biba{ri-ba ri-ba djé} ta bunita
    the hill-PL that my friend ASP live{on-them} be beautiful

A second, rather minor, relativization strategy involves the movement of the relativized PP as a whole to sentence-initial position. If the relativized noun is human, the Wh form is /kende/; if the relativized noun is not human, the Wh form is /kwa/:
(43) e kas den ku bo ta drumi ta di mi tio
do in that in which you ASP sleep be of my uncle
'the house in which you are sleeping is my uncle's'

(44) e homber ku ku
ku kende bo ta papia ta mi amigu
the man with that you ASP speak be my friend
'the man with who you are speaking is my friend'

The ungrammaticality of (45) and (46) show that this type of strategy only occurs
with prepositional phrases:

(45) xe homber kende ta papia ei ta mi amigu
the man who ASP speak there be my friend
'the man who is speaking there is my friend'

(46) xe buki kwa bo ta lesa ta bunita
the book which you ASP read be beautiful
'the book which you are reading is beautiful'

Most Papiamentu speakers seem to prefer the strategy which does not involve
movement of the prepositional phrase to sentence-initial position.

There are two curious exceptions to the account given so far of relativized PP's:
the prepositions /na/ 'in' and /te/ 'until'. These prepositions can never be
stranded:

(47a) e ora te ku nos a warda ...
the hour till that we ASP wait
'the hour until which we waited'
   b) xe ora ku nos a warda té
   c) xte é
   d) xte djé

(48a) e kas na kwa nos ta biba ...
the house in which we ASP live
'the house in which we are living'
   b) xe kas ku nos ta biba né
   c) xna é
   d) xna djé

An additional peculiarity of the preposition /te/ is that it does not occur with
the Wh form /kwa/ in the fronted position, but rather with in variable /ku/.

There is a set of rather complex late morphological rules affecting the P+PRO
combinations. We find the following sets of data:

(49) bau di N bau di djé abau 'under'
tras di N tras di djé abau 'behind'
riba di N riba djé ariba 'on, above'
ku N kuné 'with'
den di N den djé aden 'inside of'
We will assume that the forms in the third column, /abau/, /ariba/, and /aden/, result from a lexically governed PRO procliticization rule (suggestion due to Henk van Riemsdijk), which operates roughly as follows:

\[ P \quad PRO \quad PRO+P \]

As a result of this rule we find (51a) alongside of (51b):

(51a) e kas ku m'a biba den dje ...
    the house that I-ASP live in PRO
    'the house that I lived in'

b) e kas ku m'a biba aden ...

c) te kas ku m'a biba den ...

More puzzling is the contraction of /di/ + /e/, as seen in the second column. Several possibilities come to mind. A first one is that there are two rules operating: (a) a general /di/ + /e/ contraction rule which provides the input for a purely phonological gliding rule; (b) a lexically governed and sometimes optional /di/ reduplication rule.

A second possibility is that there is a general /di/ copying rule which converts \( P - di - PRO \) into \( P - di - di+PRO \), automatically providing the input for a phonological gliding rule. This rule would be followed by a lexically governed and sometimes optional /di/ deletion rule.

A third one is that there is a lexically governed and sometimes optional /di/ insertion rule, which changes \( [ \cdot \cdot \cdot \ PRO]_p \) into \( [ \cdot \cdot \cdot \ di+PRO]_p \). Considering the case of /sin e/ becoming /sindje/, the latter solution may be the most plausible one.

In the case of the preposition /ku/ 'with', which combines with PRO to /kuné/, we may postulate an underlying nasal, which only appears when the preposition combines with PRO. Given the fact that both Portuguese and Spanish have a nasal in the corresponding preposition, this would seem a plausible claim, were it not for the fact that we also get the /kuné/ contraction in comparatives:

(52) mi ta mas grandi kuné
    I be more big than-PRO
    'I am bigger than he is'

The /ku/ in comparatives is derived indirectly from Spanish and Portuguese que, which has no nasal. Curiously enough, the complementizer /ku/ 'that', which also derives from Romance que, does not contract to /kuné/:
After this rather detailed analysis of relativized prepositional phrases, we will return to the discussion of the general characteristics of relative clauses in Papiamentu. Just like x/y question formation, relativization is a process in Papiamentu which appears to violate the constraints of Chomsky's diagnostic (b) for Wh movement:

(54) e seru ku Maria ta kere ku Wanchu a biss ku mi te biba riba djé ... 
    'the hill that Mary believes that John said that I am living on'

It appears to conform to the Complex Noun Phrase Constraint:

(55) e seru ku Wanchu a ganjabo ku mi te biba riba djé ... 
    'the hill that John lied to you that I live on'

(56) xe seru ku Wanchu a sakabo mentira ku mi te biba riba djé 
    'the hill that John told you a lie that I live on'

Sentence (55) is similar to (54), while (56) is ungrammatical because /e seru/ is related to /e/ across a complex noun phrase.

As to the Wh-islands constraints, the situation is slightly more complicated, since Papiamentu admits /ku/ sentential complements where English would require a Wh complement:

(57) e mucha muher ku Wanchu a puntra su mes ku e lo laga drenta ... 
    'the girl that John wondered whether he'd let in'

The same sentence with a Wh complement introduced by /si/ is only grammatical when a resumptive pronoun is included:

(58) e mucha muher ku Maria a puntra su mes si Wanchu a lagé drenta ... 
    'the girl that Mary wondered whether John had let her enter'

In the case of (58) that the PRO in the embedded clause is not the result of a possible movement, since we have seen that only in the case of relativized PP's PRO's remain. In (59), this is not clear, of course:

(59) e mucha muher ku Maria a puntra su mes si Wanchu a kumpra flor pé ... 
    'the girl that Mary wondered whether John had bought flowers for'

Chomsky's analysis of Wh movement (1976) includes the stipulation that Wh movement "leaves a gap". We have seen that in the case of Papiamentu, elements in prepositional phrases can be relativized and questioned by leaving a preposition and a pronominal element /e/. This leads us to the curious position of having to claim that two separate strategies are operant in Papiamentu: Wh movement with NP's, no movement with PP's, while otherwise PP's and NP's are subject to much the same conditions as far as their extractability is concerned.
On the face of it, this would contradict Chomsky's remark that:

For example, in languages where relativization involves no movement rule at all but simply interprets a base-generated pronoun in the relative clause, relativization can violate the usual constraints fairly freely, as noted by Ross (1967) and many others since. A natural approach, I think, is to assume that pronouns are base-generated and permitted to refer freely (Dougherty's "anaporn relation"; cf. note 12).

In note 12, Chomsky goes on to describe the essential characteristics of the "anaporn relation", namely that:

in positions where nouns and non-anaphoric pronouns can freely occur, pronouns that can be understood anaphorically can also be understood non-anaphorically.

In fact the conditions postulated for the anaporn relation by Dougherty are not satisfied in the case of Papiamentu, since the pronominal element remains unmarked for number in the case of relativized and questioned PP's, while ordinary anaphoric pronouns are marked for number. Consider the following cases:

(60) Maria ku Wanchu a kana na kayu; m'a mira nuy uyora
Mary and John ASP walk in street; I-ASP see them yesterday
'Mary and John walked in the street; I saw them yesterday'

(61) Maria i Linda ta e muhernen ku m'a kumpira flor pe
Mary and Linda be the woman-PL that I-ASP buy flower for-PRO
'Mary and Linda are the women that I bought flowers for'

In (60) the pronoun nan clearly has an anaporn relation to Maria ku Wanchu; it is understood either anaphorically, or non-anaphorically. In (61), however, the presumed pronominal element in pe (from /pa/ + /e/), can only be understood as anaphorical with muhernen. If it were to have an anaporn relation with its antecedent, it would have to be marked for plural.

Similarly the following two sentences:

(62) e serunan ta masha lew; mi no por mira nan
the hill-PL be too far; I not can see them
'the hills are too far away; I can't see them'

(63) e serunan ku m'a drumi riba dje ...
the hill-PL that I-ASP sleep on PRO
'the hills that I slept on'

If we do not accept the hypothesis that /e/ pronouns are necessarily base-generated in PP contexts, when the prepositional phrase is relativized, and conclude that movement is involved in the same way as it appears to be with relativized noun phrases, then a specific rule such as (64) would be needed:

(64) \[
[PP \ P \ [NP \ \ ] ] \implies \[PP \ P \ [NP \ \ PRO \ ] \]
\]

This rule automatically inserts a PRO with the proper feature specifications into a PP which contains an empty NP.

If the solution sketched, involving a rule such as (64), is correct, then we can provide a unified account for question formation and relativization of
both NP's and PI's in Papiamentu. As far as relativization is concerned, we may conclude that it appears to obey the same constraints as question formation, pending a more detailed investigation of relativization into Complex Noun Phrases and Wh-islands. The problem of resumptive pronouns will be taken up again later.

**ta-fronting in Papiamentu**

In Papiamentu there is a construction like the "inverted sentence type" described by Bailey, in which a constituent is clefted, and appears after /ta/, often sentence-initially:

(65a)  
\[
\begin{align*}
\text{m'a dunabo e buki} & \quad \text{NP} \\
& \quad \text{I-ASP give-you the book} \\
& \quad \text{I gave you the book}'
\end{align*}
\]

b)  
\[
\begin{align*}
ta mi ____ a dunabo e buki & \quad \text{it's me who gave you the book}'
\end{align*}
\]

c)  
\[
\begin{align*}
ta e buki m'a dunabo ____ & \quad \text{it's the book that I gave you}'
\end{align*}
\]

d)  
\[
\begin{align*}
ta bo m'a duna ____ e buki & \quad \text{it's you that I gave the book}'
\end{align*}
\]

(66a)  
\[
\begin{align*}
m'a drumi na kas & \quad \text{FP} \\
& \quad \text{I-ASP sleep in house} \\
& \quad \text{I slept at home}'
\end{align*}
\]

b)  
\[
\begin{align*}
ta na kas m'a drumi ____ & \quad \text{it's at home that I slept}'
\end{align*}
\]

(67a)  
\[
\begin{align*}
bo ta branku; bo ta karinjoso & \quad \text{AP} \\
& \quad \text{you be white; you be nice} \\
& \quad \text{you are white; you are nice}'
\end{align*}
\]

b)  
\[
\begin{align*}
si ta branku bo ta ____ , ta karinjoso bo ta ____ & \quad \text{if it's white what you are, it's nice what you are} ' \\
& \quad \text{if be white you be, be nice you be}'
\end{align*}
\]

Presumably, these sentences have a derived structure such as (68):

(68)

\[
\begin{align*}
\text{In this section I will try to analyze the /ta/ construction within the general framework used in this paper; I will start out, however, by setting the /ta/ cleft construction apart from several superficially similar constructions.}
\end{align*}
\]
In a paper in which data from West African languages are linked with data from Afro-European Creole languages, Bynoe-Andriolo and Yillah (1975: 234-239) propose that emphatic predicate clefts in Creole languages, such as (69) and (70), should be distinguished from focussing NP clefts, such as (71) and (72):

(69) ta traha e ta traha
   be work he ASP work
   'he is really working'

(70) ta grandi e ta grandi
   be big he be big
   'he is big indeed'

(71) ta den kas di Maria b'a drumi
   be in house of Mary you-ASP sleep
   'it's in Mary's house that you've slept'

(72) ta e kachö nan a mata
   be the dog they ASP kill
   'it's the dog that they killed'

Not only are the two types of constructions pragmatically very distinct, as will be clear from the paraphrases given of (69)-(72), but also their syntactic characteristics are very different. A first difference is of course that in emphatic predicate clefts the clefted element is repeated, but not in focussing NP clefts, as is noted by Bynoe-Andriolo and Yillah:

(73a) ta kome bo a kaba di kome
    be eat you ASP finish of eat
    'that was some eating you just finished'

b) *ta kome bo a kaba di __

(74a) *ta e buki bo ta lesa e buki
    be the book you ASP read the book
    'it's the book that you are reading'

b) ta e buki bo ta lesa __
    'it's the book that you are reading'

A second difference is that NP's etc. can be clefted out of embedded clauses, under certain conditions (cf. Chomsky's diagnostic (b)), as in (75) and (76), while predicates cannot:

(75) ta Wanchu b'a bisa ku __ a bai kas
    be John you-ASP say that ASP go home
    'it's John that you said has gone home'

(76) ta e buki b'a bisa (ku) Wanchu ta lesa __
    be the book you-ASP say (that) John ASP read
    'it's the book that you said John is reading'

But:

(77) *ta lesa b'a bisa (ku) Wanchu ta lesa e buki
    be read you-ASP say (that) John ASP read the book
    'it certainly is some reading that you said that John is doing with the book'

Thirdly, we find that the constituent fronted in an NP focussing cleft can be quite complex internally (see (78)), but not in an emphatic predicate cleft, as will be seen below.
Bynoe-Andriolo and Yillah argue that the predicate element in emphatic clefts is nominalized, since it cannot appear with aspect markers:

(79a) ta traha e a traha
be work he ASP work
'he certainly did some work!' 

b) *ta a traha e a traha

(80a) iz wok i wokin
is work he working

b) *iz wokin i wokin
Trinidadian Creole

They argue for the specific NP character of the clefted predicate constituent on the basis of the fact that the primary function of the "introducer element" (Papiamentu /ta/, Trinidadian /iz/) is "to serve as a Noun Phrase Introducer" in ordinary unclefted sentences:

(81) e ta mi tio
he be my uncle
'he is my uncle'

We also find, however, PP's and AP's introduced by it:

(82) mi tio ta di Korsow
my uncle be of Curacao
'my uncle is from Curacao'

(83) mi tio ta mas grandi ku Wanchu
my uncle be more tall than John
'my uncle is taller than John'

The major difficulty with the nominalization analysis is that it obscures the differences between the focussing NP cleft construction and the emphatic predicate cleft construction, since in both cases we would have a derived structure such as:

(84) 

The interpretation rules would not be able to differentiate the two constructions in any obvious way.

The type of nominalization involved (syntactic vs. lexical, deverbal vs. deadjectival) is not at all clear. Somehow there must be a filter mechanism which has to recognize the two recurring lexical items as identical, and obviously the rightmost of the pair of items cannot be nominalized.

An alternative to the nominalization analysis which has the same empirical consequences in excluding aspect markers from the clefted predicate can take either of two paths:
(a) We postulate that the aspectual markers originate in an Aux node, immediately to the left of the VP. The base rules would then be:

(85)  \[ S \rightarrow NP \ Aux \ VP \]

\[ VP \rightarrow V \ NP \ NP \ldots \]

Automatically, then, the aspect particles cannot occur with the V in the cleft position, since that would mean that parts of two major constituents are cleft-ed together, a consequence which the grammar would have to exclude for independent reasons.

(b) We assume that the clefted element has to be bare, i.e. without any bars (cf. Chomsky, 1970). Then the clefted position would be marked \(^{+V, 0 \ bar}\). This specification would automatically exclude the aspect markers from the clefted position even if the VP expansion node were to contain a rule such as:

(86)  \[ V \rightarrow Asp \ V \]

Of course, the specification of the predicate cleft position would also exclude the specifiers of adjectives to be part of the clefted predicate constituent. This consequence will not be investigated here.

In addition to ta-fronting, there appears to be another type of fronting process: object fronting. It involves cases such as:

(87a)  m'a mira Wanchu

'I saw John'

b) Wanchu m'a mira ___

'John I saw'

There is an important difference between object fronting rules and ta-fronting; the latter cannot occur in embedded sentences, the former cannot:

(88)  e sa ku ta e buki di mi bo ta lesa

'he knows that it's my book you are reading'

(89)  *e sa ku e buki di mi bo ta lesa

The grammaticality of (88), as compared with the lack of grammaticality of (89), demonstrates the impossibility of object fronting in embedded clauses.

There are two ways of accounting for the lack of grammaticality of (89). First of all we could assume that object fronting and ta-fronting are really the same process but that there is an optional ta-deletion rule, with fairly strict conditions imposed on it:

(90)  \[ ta \rightarrow \emptyset / \# \]\[NP \]

where \# refers to the very first position in the highest clause. The existence of this rule would account for the fact that only NP's can be fronted in the described way, not PP's. It also would explain why (87b) is grammatical and (89) is not.

The other alternative involves embedded clauses in a more principled fashion. Assuming, for the sake of the argument, Emonds' (1976) framework distinguishing between root- and structure-preserving transformations, it would follow that object-fronting, which looks roughly like (91), can only apply when the COMP
is empty, i.e. in the main clause:

(91) \[ \text{COMP NP}_1 \ V \ NP_2 \ X \rightarrow \ NP_2 \ NP_1 \ V \ \_ \ X \]

Similarly, \textit{ta}-fronting would be considered a structure-preserving transformation, i.e. if any real movement occurs at all, the fronted element enters an already existing but empty position.

A crucial difference between the two types of fronting would be, under the second analysis, that \textit{ta}-fronting may take place across \textit{S}-boundaries, while object fronting typically is limited to the highest \textit{S}. The facts here are not conclusive, since intuitions on (92) vary:

(92) \textit{Wanchu Maria a bisa ku m'a mira} \\
John Mary ASP say that I-ASP see 'John Mary said that I saw'

(93) \textit{ta Wanchu Maria a bisa ku m'a mira} \\
be John Mary ASP say that I-ASP see 'it's John that Mary said I saw'

If we interpret the difference in acceptability between (92) and (93) as conclusive, the second analysis would find support. If we don't, the first analysis, involving a late \textit{ta} deletion rule, may seem preferable.

We noted already that \textit{ta}-fronting constitutes an apparent violation of the constraints of diagnostic (b) for Wh movement. One of the most obvious ways in which \textit{ta}-fronting differs from \textit{x/y} question formation and relativization in Papiamentu is the fronting of PP's. Compare the following three cases:

(94) \textit{e kas ku bo ta biba den' djé} ... \\
'the house you are living in'

(95) \textit{kwa kas bo ta biba den' djé} \\
'which house are you living in?'

(96a) \textit{ta kas bo ta biba den' djé} \\
'b) ta den kas bo ta biba' \\
'it's in the house that you are living'

It appears, however, that preposition stranding is possible when the constituent is fronted out of a deeply embedded clause:

(97) \textit{t'a mucha muher ey Maria ta kere Wanchu a bisa m'a kumpa flor pê} \\
be-the child woman there Mary ASP believe John ASP say I-ASP buy flower for-PRO \\
'it's that girl Mary believes that John has said that I bought flowers for'

Even here the equivalent with the fronted preposition appears to be preferred.

\textit{V}y\textit{i} the Complex Noun Phrase Constraint and the Wh-islands constraints the intuitions available with regard to \textit{ta}-fronting were not reliable enough nor definite enough to form the basis for a comparison between \textit{ta}-fronting and the other fronting processes in the language on this point. In the final section of this paper, the issue of \textit{ta}-fronting will be taken up again.
The NP + V strategy

The movement rules of Papiamentu generally do not affect the NP + V, subject-verb, configuration in any way. Thus verbs can only appear in emphatic predicate clefts when an identical copy appears to the right of the subject NP. If we take the analysis given in the previous chapter of the focussing NP cleft, i.e. something like (98), then it appears that the constraint on the disruption, either by movement out of it, or by movement into it, of the NP + V configuration must be formulated as a surface strategy.

(98) \[ \text{ta NP} \left[ \frac{s}{\ldots} \ldots \right] \]

The subject can be moved out of the embedded clause as long as no element intervenes between its new position and the verb (disregarding negation, aspect and some time adverbs):

-(99a) \[ \text{mi ta lesa e buki} \]
     'I am reading the book'

-(99b) \[ \text{ta mi} \left[ \frac{s}{\ldots} \ldots \text{ta lesa e buki} \right] \]
     'it's me reading the book'

The assertion that the constraint on the NP + V configuration should be formulated as a surface strategy poses some interesting questions about the relation between movement rules and morphological information available. It has been noted many times that an inverse correlation generally holds true between the amount of morphological detail that a language possesses, and its freedom of word order. If an analysis such as the one presented here could be generalized, then the lack of freedom in word order in some aspects of the grammar of Creole languages could be formulated as surface strategies.

This conclusion is particularly interesting because it provides evidence about the way Papiamentu might be related to the Ibero-Romance languages. In a very interesting paper, Klima (1970) tries to establish a relationship between the way grammars are learned and the way they are structured (cf. also Bever & Langendoen, 1972). Klima suggests that this relationship can be expressed by the following principle:

When there are multiple occurrences of the same category in one construction, without lexical or morphological differentiation, than a simple algorithm exists for distinguishing their function and no transformation will have such an effect as to interfere with the effectiveness of the algorithm.

The simple algorithm which mechanically interprets certain otherwise unspecified verb sequences without taking their transformational history into account is, according to Klima (1970), who follows Bever in this, one of the primary mechanisms by which children learn their language.

I submit that it also figures predominantly in the formation of creoles. Let me argue this briefly. Algorithms such as the interpretation of NP + V as subject + verb are used when no other (often morphological) information is present. Thus in the interpretation of sentences the sequel to such an algorithm based on surface configurations would be an interpretation of sequences using data from verbal concord and nominal agreement to trace the transformational history of the sentence. In the Creole languages, the morphological information is not present, and the algorithm still prevails. The data from
ta-fronted subject NP's reveal that it is not merely a matter of constraining
the base rules and of prohibiting movement rules involving elements of the
NP + V configuration. Of course, these constraints do occur.

In Spanish and Portuguese, a base rule generates a number of preverbal clitic
pronoun positions; these clitics can refer to direct and indirect objects,
and thus these languages do not fall within the NP + V algorithm sketched.
In Papiamentu, there is one postverbal object clitic position, which does not
create any difficulties for the NP + V algorithm. Similarly, in Spanish subjects
can move freely across the verb, which is not possible in Papiamentu:

(100) lo haremos nosotros
       'we will do it'

(101) *lo e faz nos
       ASP PRO do we

If one accepts the hypothesis that Papiamentu presents in part the result of
tries to construct a grammar of Spanish (or Portuguese), then these facts,
as well as the facts of ta fronting sketched in (99), can be explained coherently:
an algorithm was adopted that admitted into the grammar only those sentences
that conformed to the NP + V configuration, regardless of their transformational history.

There is one exception to the NP + V strategy formulated above, which at the
same time constitutes an exception to a surface filter proposed in Chomsky &
Leśnik (1977); I use here the numbering of the source quoted:

(68) *that [NP ---], except in the context [NP NP ---]

(71) The filter (68) is valid for all languages that do not have a rule of
subject pronoun deletion, and only these.

Note the following sets of data, given here for x/y questions, but equally
valid for relative clauses and ta-fronting constructions. Subject pronouns
cannot be deleted, but still Papiamentu shows instances of that NP ---:

(102a) e ta parse mi tata
       he ASP look like my father
       'he looks like my father'

(103a) ken bo ta kere ku --- ta parse mi tata
       who you ASP believe that ASP look like my father
       'who do you believe looks like my father?'

(103b) x
       And also:

(104a) e-l-a b'ai fiesta
       he-ASP go party
       'he went to the party'

(105a) ken b'a hisa ku --- a b'ai fiesta
       who you-ASP say that ASP go party
       'who did you say went to the party?'

b) x
There is no evidence for a late /ku/ contraction rule:

(106a) mi sa ku e a bai fiesta
I know that he ASP go party
'I know that he went to the party'

b) mi sa ku a bai fiesta

In English the filter is often circumvented through that deletion:

(107) Who do you think that came?  Who do you think came?

In fact, Papiamentu has a /ku/ deletion transformation, which often applies in sentential complements of verbs of saying and thinking, but the conditions for its application appear to be stylistic and essentially unrelated to the filter under consideration. These conditions merit more careful study, but have no bearing on the issue at hand.

In the face of the evidence presented in (102)-(105), we may tentatively conclude that indeed Papiamentu constitutes a counter-example to Chomsky & Lasnik's (1977) filter (68). There may be independent evidence, of course, for postulating a rule which deletes traces in Papiamentu in specified contexts before the filter applies, thus making it vacuous here. This can only be done, however, when a clear perspective has been gained relating to the extraction from prepositional phrases, and the PRO insertion rule supposedly operant there.

Conclusion

This paper was inspired by the observation that in Papiamentu both sets (a) and (b) are possible in the case of the following examples:

(108a) ta ken ta bai kas
be who ASP go home
'who is going home?'

b) ken ta bai kas

(109a) ta unda b'a drumi
be where you-ASP sleep
'where did you sleep?'

b) unda b'a drumi

(110a) ta kiko Wanchu tin
be what John have
'what does John have?'

b) kiko Wanchu tin

We will assume in the case of set (a), that ta-fronting has occurred. The question then is: how does set (a) relate to set (b)?

There is evidence that other Creole languages shared or share construction (a) with Papiamentu. Thus we find in Negerhollands:

(111) da wie bin daano
who be there
'who's there?'
And in Jamaican Creole:

(112) a huu put i de
who put it there'
'who but it there'

In Jamaican Creole /a/ also functions as the copula in certain equatives.

Bailey (1966: 90) postulates an optional /a/-deletion transformation, which converts (112) to (113), probably under the influence of English:

(113) huu put i de

Quite similarly, a first hypothesis which can be brought forward in the case of Papiamentu would state that the (b) cases of (108) (110) are the result of an optional rule of ta-deletion in the context [+Wh]:

(114) ta \rightarrow \emptyset / \text{[+Wh]}

Under this hypothesis, the fronting of Wh question words such /ken/ 'who', /kiko/ 'what', etc. is part of the more general process of ta fronting. There is one serious difficulty with this proposal, however.

If we make the reasonable assumption, as do Chomsky & Lasnik (1977), that minor deletion rules (of which (114) would be an example) are ordered after the semantic interpretation rules, then we might expect that the pairs (a) and (b) of (108)–(110) are synonymous. In fact, they are not. Note, for instance, the following pair:

(115a) kiko b’a trese pa mi
what you-ASP bring for me
'what did you bring me?'

(115b) ta kiko b’a trese pa mi

In (a), the person asking the question does not know whether the other has brought him or her something, and in (b) he or she does know that the other has something, but does not know what.

Pairs such as that of (115) suggest that the ta-deletion hypothesis is wrong for Papiamentu. An alternative hypothesis is that ta-fronting and question word fronting are separate processes, and that in sentences such as (115b), both occur; we may assume a tree configuration such as (116):

\[
\begin{array}{c}
\text{FOC} \\
\text{ta} \\
\end{array}
\quad \begin{array}{c}
\text{S} \\
\text{COMP} \\
\end{array}
\quad \begin{array}{c}
\text{S} \\
\text{kiko} \\
\end{array}
\]

If indeed the analysis presented schematically in (116) is correct, this may have some interesting consequences for our conception of the process of creolization. Why this is so will be clear after a brief survey of some other Creoles and of some African languages.
In an interesting and suggestive study Epée compares three syntactic processes in Duala (Epée, 1975: 210-226): clefting, Wh-question formation and relative clause formation, and claims that in fact the three processes constitute a single syntactic phenomenon: leftward movement of a constituent to sentence-initial position. The distinguishing feature of this leftward movement process is that, whenever a constituent is moved past the verb, the morpheme /no/ remains as a trace immediately to the right of the verb. The leftward movement process is optional.

(117a) na mende tilea wa
'I will write you'

b) wa nde na mende (no) tilea
'it's you I will write to'

(118a) o wu njika buna
'you return what day?'

b) njika buna o wu (no)
'on what day did you return?'

(119a) Kuo en muna
'Kuo saw the child'

b) muna Kuo en (no) ...
'the child Kuo saw'

While there are some differences between the three constructions: the presence of the marker /nde/ in clefts but not in x/y questions and relative clauses, the obligatoriness of fronting in relative clauses but not in x/y questions, the /no/ marker provides a diagnostic which can be used to treat a number of seemingly disparate phenomena in a unified manner.

In no way, of course, a direct link is claimed between Duala and Papiamentu. Some of the phenomena described by Epée appear to be characteristic of West African languages as a group, and may provide a background against which we can understand the Creole cases. Specifically, it appears to be the case that x/y question word fronting is a special kind of clefting:

A pattern of topicalization often appears in questions, .... Topicalization is semantically natural to questions; when we say 'It is rice he bought.', we are contrasting 'rice' with all other possibilities, and in the same way, when we ask 'What did he buy?', we are asking 'what?' in contrast with all other possibilities. (Welmers, 1973: 416/7)

I don't know of any study which deals with topicalization in West African languages in sufficient detail to make a comparison possible between African topicalization patterns and Western European Wh movement patterns. They may well have very different characteristics, and the generalization made here over the African cases may well be illegitimate.

In any case, 'topicalization' (or 'focussing' or 'clefting', the terminology appearing in the literature is rather confusing) and 'Wh movement' have in common that both involve the fronting of specific elements. I submit that the Creole cases, such as the Papiamentu one described here, present a reanalysis of the European data in terms of African patterns. It is this reanalysis, on the basis, to be sure, of superficial similarities between the two cases, which has led to the contrast between sets (a) and (b) of (108)-(110). Only when we know more about specific characteristics of the African fronting
transformation will we be able to specify the form which this reanalysis took.

There is some slight evidence that in early forms of Creole languages x/y questions could be formed without movement being involved. Thus we have the following example from Schuchardt's description of Príncipe Portuguese Creole, which may represent one of the most conservative Creoles known:

(120) ninghe' ti sa fala koli a
someone you ASP speak with-PRO Q
'with who are you speaking?' (Schuchardt, 1889)

It may be possible to analyze (120) as a case in which an indefinite element is generated in sentence-initial position, and a PRO element referring to it within the clause itself. Alternatively, we may analyze (120) in a similar way as the Papiamentu examples: involving Wh movement and PRO insertion.

Ferraz (1974: 137, 190) discusses the PRO elements which remain in Sao Tomé Portuguese Creole in relativized and questioned prepositional phrases, and claims that they result from an African substratum. Thus, he claims that Zulu (121) or a different example from another West African language can be seen as having provided a substratum for Sao Tomé Creole (122):

(121) indu umuntu alala kujona ...
house person live in-PRO
'the house in which the person lives'

(122) .'ka mú ku nga 'vive ne
house me that I-KA live in-PRO
'the house that I live in'

Whereas modern Standard Portuguese does not permit resumptive pronouns in relative clauses, non-standard forms of the language evidence an abundance of examples of constructions such as (121)-(122):

(123) esse e o cavalo que vou nele
that is the horse that go-1s on-PRO
'that is the horse I go on'

Also in other non-standard varieties of the Romance languages we find equivalent examples.

These cases are different from the Papiamentu examples in that the resumptive pronoun in the relative clause agrees with the antecedent in gender and number, while in Papiamentu no agreement occurs. On the basis of this fact we have argued that Papiamentu possesses a PRO insertion rule. This PRO insertion rule may be the result of another process in which African structures involving base-generated resumptive pronouns were mapped onto European structures, involving gaps resulting from Wh-moved constituents. Here again, more information will be needed to determine the precise nature of this mapping process.
Notes:

1 This paper is a much revised and expanded version of a chapter of my M.A. Thesis (Universiteit van Amsterdam, 1974). I am grateful to Raúl G. Römer and Simon C. Dik for their comments on that earlier version, to Linda Richardson and Murella Roberta for providing and interpreting the Papiamentu data and for their support, and to Henk van Riemsdijk and particularly Hans den Besten for their help in making sense of the material and in providing a framework in which to interpret it. Of course, all mistakes in fact and in interpretation which remain are mine.