PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/14650

Please be advised that this information was generated on 2019-12-26 and may be subject to change.
THE SEARCH FOR UNIVERSALS IN LANGUAGE GENESIS: ÉTAT DE LA QUESTION AND RESEARCH PROGRAM*

Pieter Muysken
Universiteit van Amsterdam

0. Introduction

There is a large literature on universals in language genesis, which I will not try to summarize here, referring the reader to such recent introductions to the field as Holm (1988), Mühlhäusler (1986), and particularly Romaine (1988), and to Mufwene (1990). A picture both rich and confused emerges there, of which the following is but a gross simplification.

This paper briefly examines two dominant approaches to the issue of universals in language genesis, and proposes a third one, which is then fleshed out somewhat in a sketch for a research program. The proposal focuses on the need to develop a more explicit universal theory of language contact, in which pidgin and creole data are combined with those from other fields in language contact studies: borrowing (lexical, phonological, and grammatical), intercultural communication, second language learning, code-switching, etc.

1. Procedural and constitutive universals

The two dominant conceptions of universals in the genesis of languages can be termed procedural and constitutive. To understand what these two conceptions are about, consider the very schematic figure in (1), depicting the circumstances under which languages can emerge (again I refer the reader to the introductions listed above).

(1) CONTACT: PROCESS X ---» PIDGIN

\[\downarrow\]

PROCESS Y ---» CREOLE

Now the nature of the processes involved in genesis is quite controversial. Some researchers deny that all creoles had an identifiable pidgin ancestor, and indeed there is a great deal of disagreement on what pidges may look like: are they very unstable and limited systems, or can they be conventionalized and relatively expanded. Finally of course, there is much disagreement on the nature of the processes X and Y.

The reason that universals are mentioned at all when we are discussing language genesis is that the newly formed languages are often assumed to be alike in ways that cannot be explained exclusively through reference to the similarities between the contributing languages. Hence the appeal to universals, and the universals involved are the two types mentioned: procedural universals, universal properties of processes X and Y, and constitutive universals, universal properties of the resulting pidges and creoles.
1.1. Procedural universals

Procedural universals generally have a (more or less precisely defined) psycholinguistic basis and are formulated as strategies a speaker may employ in the language contact situation.

For pidgin genesis universal properties are attributed either to the adaptation mechanisms of speakers of the dominated languages, as possibly in (2), or of the dominant colonial languages.

(2) Disregard pre-verbal unstressed elements in the target (Schuchardt 1883: 237)

While (2) can account for the fact that the Romance pre-verbal clitics have disappeared in the creole languages. Since the strategy is one of reduction, it is assumed to characterize pidgin genesis. It is problematic in that it cannot correspond to a learning process in the form formulated. If we assume it is part of one of Slobin's operating principles (e.g. 1978) governing language learning, as in (3), this problem disappears:

(3) Pay attention to the end of the word.

Anthony Naro (1973 and later work) has formulated principle (4) to describe the way speakers of the dominant languages may have adapted their speech in the language contact situation:

(4) Express each separately intuited element of meaning by a phonologically separate stress-bearing form.

Such a principle would explain the amount of semantic transparency characteristic of some aspects of pidgin and creole systems.

For creole genesis we must think of the kind of developmental universals suggested (but never formulated clearly) in the work of Mühlhäusler (1986), governing the gradual expansion and development of systems of signifiers.

I have no quarrel as such with procedural universals. Undoubtedly there are general mechanisms of language learning. I do find them rather vague in many cases.

1.2. Constitutive universals

Constitutive universals belong to the domain of some theory of grammar. A first instance is (5):

(5) Every natural language must conform to Universal Grammar.
This principle, which has its basis clearly within the generative research tradition, is interesting in two respects. First of all, it undercuts the basis for the commonly held assumption that creoles are in some sense special languages.

Second, it potentially allows us to make a principled grammatical distinction between pidgins - need not conform to UG - and creoles - must conform to UG. This in turn could lead to research into the development from pidgin to creole where we could see UG in action, as it were. The trouble is that strong versions of the Interlanguage Hypothesis, which holds that even intermediary products of second language learning are natural languages, would imply that pidgins, inasmuch as they are like interlanguages, fall under (5).

Thus (5) offers not much of particular interest, however true. A principle such as (6) is more promising:

\[(6) \text{ Creole languages present the unmarked option for each parameter of Universal Grammar.}\]

This proposal was made concretely by Bickerton (1984), but the idea behind it appears in much work from early on in the field of creole studies. If it were plausible, it would lead a fruitful set of research questions: we would have an independent grip on markedness theory. To get some idea of what features creoles have in common, consider (7) and (8):

\[(7) \text{ wanpela man i bin skulim mi long Tok Pisim} \quad \text{TOK PISIN}\]
\[
\text{ one man PR ANT teach me in Tok Pisin} \\
\text{ 'A man was teaching me Tok Pisin'}
\]

\[(8) \text{ so mo ka ta toka palmu} \quad \text{SENEGAL CRIOULO}\]
\[
\text{ one hand NEG HAB touch palm} \\
\text{ 'One hand can't touch its palm'}
\]

Generally we find a subject-verb-object word order, and between the subject and the verb different particles can occur. In the Tok Pisin example (7) a predicate marker i and an anterior tense marker bin. In the Sénégal Crioulo example (8) these are a negation marker ka and a habitual or generic marker ta. Often these particles have a relatively fixed order, roughly the one in (9):

\[(9) \text{ negation (predicate marker) TENSE MOOD ASPECT}\]

In addition, many of the semantic categories expressed by these particles occur in creole after creole as well. Anterior tense (marking a moment in time prior to a reference point) is more common than absolute past (marking a moment in time prior to the moment of speech) tense. Often there is an irrealis mood category, which expresses counterfactuality, distant future, volition.
SEARCH FOR UNIVERSALS

There are several problems with (6), however. First of all, the colonial history that gave rise to most creoles imposes serious limitations on the data-base. The colonial languages are all Romance or Germanic, and many of the African languages involved in the Atlantic creoles are Niger-Congo, or even more restrictively, Kwa. Thus, it is hard to see whether many of the features that creoles have in common are not simply due to common ancestry.

Second, creoles do not present a unified group when we consider some of the more familiar parameters that have played a role in research in the past decade, such as pro-drop and preposition stranding. Many creole languages show the pattern of Papiamentu in (10):

(10) a. e ta kome
    he ASP eat
    'He is eating'
    (compare Spanish él está comiendo)

b. *[0] ta kome
   (compare Spanish está comiendo)

Notice, however, that with impersonal verbs, there are differences:

(11) a. [0] ta parse ku Wanchu a bini bek
     'it seems that John has come back'

b. [0] semble Jan pati
   'it seems that John has left'
   (Deprez, 1990)

c. a gersi dati Johnny doro
   'it seems that John has arrived'

Here the pattern of the creoles does not follow that of the three colonial lexifier languages (Ibero-Romance pro-drop, French and English no pro-drop) completely for Haitian. Moreover, there are cases of English-related creoles, so far little studied, with pro-drop as well, such as Miskito Coast English Creole (O'Neill, ms.).

In addition, we find patterns of absent subjects in Pacific creoles such as Hawaiian English:

(12) sam gaiz samtaimz [0] dei kam
    'Sometimes some guys come'
(13) difren bilifs [0] dei get, sam gaiz
    'Some guys have different beliefs'
    (Bickerton, 1981)
In (12) and (13) *dei* can be argued to be a pre-verbal particle marking a subject position which can remain empty, either through left-dislocation as in (12), or through inversion, as in (13).

Admittedly, the whole discussion surrounding pro-drop has become more complex, and the original optimism concerning a single parameter has died down. Also, no creole system has been thoroughly explored yet, as far as I know, with regards to possibilities for empty subjects. Nonetheless, initial results are not promising as regards the possibility of establishing a uniform parameter setting for pro-drop for a larger group of creole languages.

Similarly for preposition stranding. Notice that even an English-based creole such as Sranan does not allow preposition-stranding (English itself allowing it of course). This would confirm the idea that not-stranding is the unmarked option, as Van Riemsdijk (1978) had assumed:

(14) * a nefi san a e koti a brede nanga ___
    the knife that he ASP cut the bread with
    'The knife that he cuts the bread with'

When we look at the creole languages more generally, three groups can be discerned, however. One small group of English-based creoles and Dutch-based creoles that allows stranding, a large group that does not allow stranding, and a group of Ibero-Romance based creoles that allows stranding only when the resulting gap or trace is realized as an unmarked pronoun.

(15) a. stranding allowed: Jamaican
    Krio
    Berbice Dutch
    Negerhollands
    b. stranding not allowed: Saramaccan
    Sranan
    Haitian
    c. stranding with trace spell-out: Papiamentu
    Capeverdean

What makes the case of Papiamentu special, and distinct from languages which allow resumptive pronouns in prepositional phrases (e.g. in relative clauses), such as Sranan, is that in Papiamentu the spelled-out traces do not show number agreement when the antecedent is plural, while resumptive pronouns do.

Again, there is considerable variation rather than the uniformity that principle (6) would have us believe. None of this comes as a surprise if we assume that creoles are just like other natural languages, distinguished through parametric variation. For me it makes attempts to establish constitutive universals for creoles less promising, however.
SEARCH FOR UNIVERSALS

This should keep no one from working on creoles just as one would on other languages: indeed they are quite complicated systems that have much to offer to grammarians. In addition, since they share some features (as well as not sharing others) with well-studied Western European languages such as English, Dutch, French, and Ibero-Romance, the type of research into subtle parametric differences between related languages that has characterized orthodox generative work in this last decade can be fruitfully pursued as well. One can imagine a series of sophisticated grammatical comparisons of Haitian and French, Jamaican and British English, etc. and indeed some of those have started appearing (DeGraff, 1992; Lefebvre et al., 1990; 1991; 1992).

2. Interactive universals: a research proposal

The question then becomes: is there any theoretical basis for studying the creole languages as a group? My answer is: yes, but only as part of a larger range of phenomena of language contact. Where creoles can and should contribute is toward a more general theory of possible interactions between grammatical systems: interactive universals. Not just the colonial languages contributed to the creoles, but also the dominated languages of Africa, Asia and the Pacific - termed substrate languages - played a crucial role in the formation of many of them as well.

The proposal I want to make goes against a train of thought characteristic of much recent work in the field, which runs roughly as follows: explaining language genesis can proceed either by postulating language universals (procedural or constitutive) or by appealing to the contribution of particular substrate languages. Thus, to phrase the contrast in terms of a great nineteenth century debate, substrate explanations could be considered historicist - in terms of accidents of human history - and non-substrate ones as romanticist - in terms of inherent properties of the systems involved. A conference held in Amsterdam a few years ago stressed this contrast in its title Universals versus Substrata in Creole Genesis (a book appeared with the same title; cf. Muysken & Smith, 1986). Universals are exciting, substrate is particularistic, and hence boring, in this perspective.

What is needed is in fact a theory of how substrata can play a role, where they do and where they do not. This theory cannot be formulated by looking simply at the evidence of the creole languages, since many of them emerged over three hundred years ago and since often the evidence is ambiguous and controversial. What we must do is through studying more contemporary cases of language mixing, see what general properties these have, and then go back to the evidence of the creoles. Here I will sketch one example of this, to show how this research could proceed. In the conclusion, I will briefly sketch some other areas where relevant research can be carried out.

It will be difficult to really deal with the issue here to what extent the substratum languages have played a role in the genesis of creoles. That question is far too generally put. There are cases where they clearly have, such as Berbice Dutch Creole (Kouwenberg 1991; Smith, Robertson, and Williamson, 1987) and Gulf of Guinea Portuguese Creole
(Ferraz, 1975, 1979), and others where evidence for substrate is very slight, such as Reunion French Creole. Neither the first set of cases nor the second one can be taken as 'typical' or 'exemplary'. Rather we should set up a theory which adequately accounts for the pattern of substrate influence where it is arguably present.

2.1. Presumed substrate properties of the Caribbean creoles

Caribbean creoles have often been assumed to share certain features of the Kwa-group of West-African languages, notably Fon and Akan (Alleyne, 1981; Boretzky, 1983). This has also been disputed, most recently in a remarkable paper by Bickerton (1989). For the purpose of my argument it is not important whether the assumption is actually correct. The point of my argument is that the presence of substrate influence is as interesting, from a universalist perspective, as its absence, and as much constrained by linguistic principles.

A first reputed feature are the serial verb patterns (Voorhoeve, 1975):

(16) e-l-a bula bay
    he ASP fly go
    'He flew away'

(17) li pote sa bay mo
    he bring that give me
    'He brought that for me'

In (16) bula 'fly' and bay 'go' (both derived from Ibero-Romance) form a verbal chain, and are predicated of the subject together. Similarly for pote 'carry' and bay 'give' (both derived from French verbs) in (17). We find many cases of serial verb chains in most Caribbean creoles. Another example is given in (18) below, where teki 'take' and koti 'cut' form a chain.

A second potential substrate feature is predicate cleft (Bynoe-Andriolo and Yillah, 1975): the copying of a verb for focus in sentence-initial position, often after a particle which may be the copula or be related to it:

(18) na teki edgar teki a nefi koti a brede
    be take Edgar take the knife cut the bread
    'Really with the knife Edgar cut the bread'

Again, most Caribbean creoles and a number of West-African languages share this feature.

A third, fairly obvious feature concerns numerous idioms (e.g. Hancock, 1980; Turner, 1949) that are very similar in the Caribbean and in West-Africa. Idioms are tricky because often the type of metaphor implied is almost universal, but still the number is too large and the similarities too great to discount it.
A fourth feature concerns ideophones, which occur in many Caribbean creoles as well as in West-African languages. Some examples from the Surinam maroon language Sarrancan, where they are particularly prominent (Rountree & Glock, 1982: 145-7):

(19) mi hopo \textit{vu}u \quad 'I got up suddenly'
mi naki \textit{en g}bo \quad 'I hit it with force'
a weti \textit{fa}an \quad 'It's really white'
a guun \textit{pi}i \quad 'It's really green'

A fifth property concerns subcategorization and selection features of verb classes (Koopman, 1986). Koopman argues that the types of complement selected by Haitian verbs closely resemble that of verbs in different West-African languages. The same point made in recent work by Lefebvre and her colleagues.

On the other hand, the creoles also lack a great many features of the West-African languages. As to word order, West-African languages are often underlyingly OV, and Caribbean creoles very consistently VO. In addition, the tense/mood/aspect systems of the two language groups do not correspond, and neither do the nominal determiner systems.

Finally, we find that morpho-syntactic categories do not correspond. An example in point are the question word systems of Fon and Saramaccan, studied by Smith (cf. Muysken & Smith, 1990). These systems are of interest here since Smith can show that the Saramaccan words \textit{ambe} 'who' and \textit{andi} 'what' derive from dialects of Gbe, closely related to Fon. There is evidence for direct transmission. Nonetheless, the semantic organization of the rest of the paradigm is not parallel:

(20) Fon: \quad \begin{tabular}{lll} 
WHO & \textit{m} -tm* & person-Q \textit{m} (fusion) \\
WHAT & e-t /ani/*nu- & that-Q/what/*thing-Q \textit{n} (fusion) \\
WHICH=A & t & Q \\
WHEN & hwe-t (-nu) & time-Q-(thing) \\
WHERE & fi-t /fi & place-Q/fi (fusion) \\
WHY & e-t /-u(tu)/ani u(tu) & that-Q-body/what-body/ \\
 & ani gbe/n -gbo & what-aim/what-towards \\
HOW & n \ldots \textit{gbo} & what \ldots by \\
\end{tabular}
While two basic forms from Fon occur in Saramaccan as well, the rest of the question word system is not modelled or calqued upon the Fon system.

Clearly, what is needed to make a plausible case for substrate influence is an account that explains the pattern of transmission found: lexical properties of content words survive, properties of function words and other grammatical patterns do not.

2.2. Relexification

I would like to argue that the pattern of relexification found in the case of Media Lengua, a half-Spanish, half-Quechua vernacular contact language from Ecuador provides such an explanation (Muysken 1981), if the process of relexification is assumed to have been responsible for the substrate influence.

Whether it does or not is a complicated issue (cf. Muysken, 1988). Ordinarily, it is assumed that processes of second language learning with a reduced input are responsible for the formation of the plantation pidgins that lead to the Caribbean creoles. However, the second language acquisition literature is fairly conclusive about the fact that, as far as is known, the type of transfer that would be needed to produce substratum effects is generally absent in second language learning. For a while it was thought that ordinary processes of second language learning would involve positive transfer from the substrate languages, but the extensive study of second language learning processes in the last two
SEARCH FOR UNIVERSALS

decades makes such an assumption unwarranted. Thus the type of second language learning that must have lead to the plantation pidgins must have been of a particular type: non-target directed, L1-based, etc, and this may be similar to relexification.

Consider first some examples. (22a) is the Media Lengua form, (22b) the corresponding Quechua, and (22c) the corresponding Spanish form.

(22) a. unu fabur-ta pidi-nga-bu bini-xu-ni MEDIA LENGUA
    one favor AC ask NOM for come PR 1
    'I come to ask a favor'
b. shuk fabur-da maña-nga-bu shamu-xu-ni QUECHUA
    c. vengo para pedir un favor SPANISH

Notice among other things that the Media Lengua version corresponds to the Quechua version morpho-syntactically. As to lexicon, the Quechua version only has one Spanish borrowing, fabur, while Media Lengua roots are all from Spanish. The same phenomenon can be observed in (23) and (24):

(23) a. kuyi-buk yirba nuwabi -shka MEDIA LENGUA
    cavia for grass there is not SD
    'There turns out to be no grass for the cavias.'
b. kuyi-buk k'iwa illa-shka QUECHUA
    c. No hay hierba para los cuyes SPANISH

(24) a. yo-ga awa -bi kay -mu -ni MEDIA LENGUA
    I TO water LO fall CIS 1
    I come after falling into the water.
b. ūnuka-ga yaku -bi urma-mu -ni QUECHUA
    c. vengo despues de caer en el agua SPANISH

What we find in Media Lengua is a process, not so much of extensive borrowing, but of large-scale relexification: Quechua lexical roots have been replaced by Spanish ones. At the same time the Quechua affixes, clitics, and grammatical structure have remained intact. Relexification needs to be invoked if we want to account for the possible survival of West-African features in the Caribbean creoles.

What is important now for the problem of substrate influence raised before is what happens with morpho-syntactic categories in the process of relexification: we saw that the West-African morpho-syntactic categories have not survived in the Caribbean creoles. Interestingly enough the same holds for Media Lengua. I have argued this elsewhere in detail for the system of deictic elements (Muysken, 1988). A simple example is also the question word system of Media Lengua:
We can conclude from (25) that the Media Lengua forms do not derive directly from Quechua semantic equivalents. Neither do they immediately correspond to the Spanish system. I will now try to relate these findings to the question posed in the introduction of how we can have both transmittance of substratum features without actual phonological shapes surviving and at the same time restructuring of those features.

What Media Lengua shows at least is that there can be processes of relexification on a large scale within a speech community, and that the process of relexification only leads to partial maintenance of L1 structures. Consider the difference between the relexification of content words and of morpho-syntactic categories, i.e. function words. In the earlier case there is straightforward relexification, generally maintaining the Quechua semantic distinctions. In the latter case, there is drastic restructuring of the system. How can we explain this difference? Relexification can only take place under semantic resemblance. When the Quechua verb *riku-* 'see' is relexified as *bi-* (from Spanish *ver*), this is possible because there is a large shared element of meaning, established through reference to some extra-linguistic mental representation.

Now take function words. These do not have a meaning outside the linguistic system that they are part of, since their meanings are paradigmatically defined within that linguistic system. So when you relexify a system of function words, automatically the semantic organization of the target language comes in, and the result is at best a compromise between source and target language systems.

This conclusion is relevant to the substrate debate in creole studies as a whole. If the argument is correct, we must conclude that the only African features that could have been transmitted more or less intact through relexification are those dependent on properties of content words. This means: lexically determined semantic distinctions and subcategorization features, but not syntactic properties related to function words. This consequence seems to me more or less on the right track, given the conflicting evidence for substratum so far. The strongest cases involve lexical properties of function words. Only content words can be relexified without restructuring, and hence grammatical elements of the source language never appear with their original properties.
SEARCH FOR UNIVERSALS

Notice that serial verbs patterns, idioms, ideophones, and selectional properties all involve lexical characteristics of content words. The same is not obviously true for predicate clefts, as in (18) above: they would seem to represent a structural rather than a lexical feature of the Caribbean creoles. One possible way of arguing that predicate clefts are lexically determined is by assuming the fronted verbal element to be a nominalized infinitive. While this is not implausible, evidence so far has not been conclusive.

3. Other possible avenues of research

My proposal is to integrate creole studies within the general field of language contact research by trying to formulate general principles governing grammar contact, that is to say, interactive universals, on the basis of the study of contemporary situations of language contact and to then apply these universals to the interpretation of the creole evidence. I have illustrated this with recent results from the study of relexification processes, but other possibilities come to mind.

One area to consider is lexical borrowing. Much recent work confirms the idea that there are hierarchies of borrowability, perhaps along the lines of (26):

(26) N - A - V - P - Adv - Neg - Conj - Pron - COMP - ...

Some elements are borrowed more easily than others. Quite apart from the question to what extent these hierarchies are universal or specific to individual language pairs, and how these hierarchies are to be accounted for, we can apply insights in this area to the study of creole genesis. Can we describe this process in terms of borrowing into nascent structures, and what does this imply for the lexical categories we are likely to encounter?

A second area concerns accommodation of surface patterns. Silverstein (1972) has originally suggested that Chinook jargon emerged through surface compromises between Chinook and English. While this conclusion has been contested there is work by e.g. Kouwenberg (1991) on Berbice Dutch which suggests that such mechanisms may be operant. We may look at contemporary studies of foreigner talk and conversational accomodation to see what mechanisms regulate such accomodation and what constraints there are on such a process.

A third area, hitherto neglected, concerns phonological dimensions of language contact, and particularly the phonology of borrowing and second language acquisition. The way that phonological systems interact is possibly quite different from the way morpho-syntactic systems interact, and a more general and consistent study of such interactive processes may well give us insight into the way creole phonological systems emerged.

What I hope to have made plausible is that there is room for a fruitful research program on interactive universals, which can help link the study of pidgins and creoles to the more general study of grammar contact.
Note

* Much revised version of a paper presented at a workshop on universals (organized by R. Zuber) at the Université Paris 7, Jussieu, June 1990. I am grateful to comments from an anonymous reviewer.

References