Semantic Transparency as a Factor in Creole Genesis

Pieter A.M. Seuren
Herman Wekker

In this paper we propose to discuss the transparency principle as one of the factors that may be said to play an important role in creole genesis. We shall deal with the notion of semantic transparency (ST) from two very different theoretical perspectives: that of formal linguistics and that of second language acquisition (SLA) research. In recent years the transparency principle has received a great deal of attention from workers in the area of child language acquisition (e.g. Slobin 1977, 1980, 1982) and historical linguistics (e.g. Lightfoot 1979), but its relevance to the study of SLA and pidginization/creolization has hardly been explored. In a preliminary attempt, Wekker (1982) showed how in one specific grammatical area (that of the expression of counterfactualness) the notion of semantic transparency can be applied to the study of both SLA and pidginization/creolization processes. In simple terms, his claim was this: second language learners and learners of a pidgin/creole language will both tend to re-analyze highly marked and opaque areas of the grammar in an attempt to make the language easier to learn, to understand or to produce (see also Kellerman 1981, 1983). However, although the notion of transparency is by no means new (cf. Martinet 1955), it has so far been insufficiently defined and constrained, and what has been lacking, in particular, is a formal unified theory of semantics, syntax and morphology to which the notion can be explicitly related. In this paper, given the limitation of time, we shall concentrate on only three important points. First, we propose to discuss the notion of semantic transparency, as seen against the background of creole genesis as well as in relation to SLA. The interesting question here, for example, is to what extent the ‘universals versus substrata’ debate resembles the current ‘universals versus transfer’ controversy in SLA research studies. Secondly, we shall provide a brief characterization of a theoretical framework that seems to be suitable for the description of linguistic change, in particular that which leads to the genesis of creole languages.
And finally, we shall (again briefly) outline a research programme that follows on from our proposals. The basic research question that we shall formulate is this: is transparency a universal operating principle which plays a role in creole genesis, as it seems to do in other developmental processes, and if so, what are its consequences? We shall, however, have space for only a few concrete examples to illustrate our ideas.

Let us begin with a few preliminary remarks. We start from the commonly accepted assumption that creole languages are languages that have developed out of a preceding pidgin. When adult speakers of a language A are forced by circumstances to communicative with speakers of a language B while the speakers concerned are not competent in the other language, and the same circumstances make it imperative that the language of communication must be B, not A, then, especially when the A-speakers are of low educational or social level, some sort of improvised lingo will develop approximating B but differing from B in many important respects, throughout the grammar, the lexicon and the phonology. Such a ‘broken lingo’ is technically called a pidgin. and pidgins, as we know, are characterized by a high measure of individual variation and instability as well as by a heavy dependency on features of context, situation, and background knowledge presumed to be available. Pidgin speakers, in other words, depend heavily on what we shall call semiotic improvisation. Semiotic improvisation is the general term used here to cover the communication strategies that learners of different proficiency levels employ in attempting to solve the problem of having to express themselves with limited linguistic means. (For a survey of the relevant SLA literature, see Poulisse et al. 1984.) Pidgin speakers, like second language learners, will resort to all sorts of expressive means, often non-linguistic ones, to express more or less what they intend to communicate. Pidgin utterances grossly underdetermine the meanings intended. They are characterized by the stringing together of lexical elements in haphazard grammatical shapes, while the listener must make heavy use of his powers of combination and inference to reconstruct the intended meaning. Semiotic improvisation is based on general cognitive strategies for symbolic expression, not on specifically linguistic strategies as developed by the human race. In a sense, therefore, one is entitled to speak of a more ‘primitive’ use of language in the case of pidgins.

Pidgin utterances are also characterized by certain strictly grammatical features. For example, there is no functional (as opposed to petrified) morphology to speak of. All semantic elements that are expressed are cast into
SEMINATIC TRANSPARENCY

separate lexical words, whose meanings are adapted, if necessary, to the expressive purpose at hand. The order of the words is strongly influenced by the native language A of the pidgin speaker (we shall speak of background language or BL, but to some extent also by the target language (TL) B. (It is, of course, also constrained by whatever logical requirements there are for word order with respect to logical properties of intended meanings, such as quantifier scope.) To give a modern example, one can often hear foreign visitors in Italy who speak no Italian asking for *uno panini* when they want a sandwich. They have not acquired the morphophonemic rule of Italian which prescribes the form *un* in this case, nor do they know that the suffix -i in *panini* indicates a plural. The Italian caterer, on his part, is not bothered by these linguistic impurities and will provide the sandwich desired. He might even condescend to use such pidginised expressions himself if he finds it commercially advantageous. Or take the Turkish speaker in Germany, who might say something like *Das 'gut' bedeutet* (i.e., "that means 'gut'") with Turkish SOV-order, or *Er das habensie nich gesagt* (i.e., "he hasn't said that") with petrified morphology, and even syntax, in *habensie*, as well as SOV-order.

Such cases of pidgin, however, are unlikely to develop into a creole. Turkish children born in Germany will grow up with a dual command of both German and Turkish if they are fully exposed at an early enough age to both languages. Creole languages arise typically when children are fully exposed to a pidgin but not (at least in most cases) to the corresponding TL, whereby the pidgin may or may not be accompanied by the BL. In other words, a pidgin turns into a creole when it becomes native. The first generation of native 'learners' of creole perform what appears to be the miracle of turning a largely improvised and highly semiotic means of communication into a real language with its own rules and grammar. The process of language acquisition is more, much more, 'creative' than in standard cases.

It is this creative achievement of the first generation of creole speakers that has confirmed linguists in the assumption that humans have an innate capacity to develop specifically human languages. The process of creolization is regarded as confirmation, in a general and global sense, of linguistic innateness. And as long as we stay at this level of generality we have no quarrel with this point of view: it is no doubt true that there is a strong and highly specific innate linguistic faculty in humans. But if we want to go beyond general philosophical questions and find out what this innate faculty specifically consists of, and how creolization can be seen to confirm the reality of this faculty, we must engage in actual research. This means that we must begin by posing, and
if possible answering, a few questions.

One important question is this: what precisely is meant by “innate”? If we say generally that whatever is universal in human language is innate, we may be saying something that is true, but such a statement is also quite unrevealing. We have to find out which universal properties are substantively built into the human organism, and which are more or less accidental by-products of the way the organism is structured and the circumstances of operation of the organism. The latter are epiphenomenal upon the organism. A few examples will illustrate this. The fact that normal body temperature is around 37 centigrade is not substantively built into the organism but is an epiphenomenon of it. Just as fever is an epiphenomenon of certain inbuilt forces that come into action under certain external conditions. The property of males, on the other hand, to grow a beard after puberty is substantively innate in that it is a purely biological property of the organism at a certain stage of development.

In the case of language, two extreme positions are in principle possible: one can hold that all linguistic universals are epiphenomenal upon general cognitive functions, or one can defend the view that they are all substantively innate. Neither position seems reasonable, however, in the overall perspective of what we know about language and cognition. Some intermediate position in this debate seems more plausible. That is, one would have to detect some ‘axiomatic’ structure in linguistic universals. Given a preferably small number of axiomatic or basic universals, which can be regarded as being substantively innate, the remaining universal linguistic features would have to be shown to be derivable from the basic universals as epiphenomena provoked by the machinery in question, with or without interaction with external circumstances. In fact, our hypothesis that the genesis of creole languages is heavily determined by a tendency to maximize ST fits in naturally with this programme of setting up basic linguistic features and deriving other features as epiphenomenal from these.

But then, given this distinction between substantively innate and epiphenomenal linguistic features, the question arises as to how precisely creolization processes can be taken as confirmation for the innateness view of linguistic universals. The question will be easy to answer if creole languages turn out to be just ordinary natural languages without any specific linguistic features characterizing them as being creole languages. If there are no specifically creole linguistic universals, then indeed one could maintain that linguistic universals are all, or virtually all, substantively innate and that first
generation creole speakers put this whole innate apparatus to work in order to create another exemplar of the species 'natural language'. Creole genesis would then be confirmation for the extreme view that all linguistic universals are substantively innate.

However, most creolists are of the opinion that there are linguistic features which typically characterize creole languages, though they are far from unanimous on what these creole linguistic universals amount to. But there is consensus on at least one point. It is generally agreed that creole languages have little or no morphology and that they place the burden of semantic expression on syntax and the lexicon. Other creole universals have been proposed, for example that all creole languages have SVO word order, that they all have verb serialization, that they all have copula deletion before adjectives, that they all have a preverbal particle system for Tense, Modality and Aspect, in that order. Yet, although the phenomena in question are no doubt highly intriguing, there are either too many clear counterexamples, or there is not enough reliable information available on the actual facts. But the absence (or extreme poverty) of morphology in creole languages seems to be a solid datum, and a highly significant one. For its known from general linguistic theory that there are universals of morphological systems. The absence of morphology from creole languages now means that not all universals have the same status. First generation creole speakers apparently activate some universals but not others.

We know that no language makes use of all linguistic universals: every language makes, so to speak, its own choices. What is striking in this case is that the group of languages that are characterized on independent (historical) grounds as creole languages have at least one, and possibly more, choices in common. If this fact is significant, as it appears to be, then it should teach us something about universals as well as about factors at work in language genesis. (It will be clear that an independent characterization of the group of creole languages is essential: one cannot say that any language that lacks morphology is therefore also a creole language.)

Given the distinctions made so far, it is reasonable to assume that substantively innate features must be activated by creolizers, by biological necessity, just as instinctive behaviour in animals is biologically necessitated. Epiphenomenal features, on the other hand, need not all occur in a creolization situation: that will depend on circumstances. Having observed that creole languages generally lack morphology, we feel entitled to conclude that whatever morphological universals there are, will be epiphenomenal, not
substantive.

While this in itself may be an interesting conclusion, we feel that there is sufficient ground for putting forward the claim that the lack of morphology is a specific consequence of a more general feature of creole languages, their property, that is, of maximizing semantic transparency. The notion of ST, as we have said, is sometimes referred to in the applied linguistics literature, yet it must be admitted that it has remained vague and essentially undefined. Langacker (1977:110) ventured the following formulation: ST is presented as an operating principle or strategy by which learners strive for an “ideal or optimal linguistic code [which] will be one in which every surface unit, typically a morpheme, will have associated with it a clear, salient, and reasonably consistent meaning or function and every semantic element in a sentence will be associated with a distinct and recognisable form”. This is not an empirical definition, since this formulation relies too much on intuitive terms such as “clear”, “salient”, “reasonably consistent”. But it contains some conditions to be fulfilled for a full application of the ST strategy. In particular, this formulation stipulates a strict one-to-one correlation between semantic elements and structural surface units (morphemes).

In similar vein, Slobin and Naro described the function of the transparency principle as a tendency “to maintain a one-to-one mapping between underlying semantic structures and surface forms, with the goal of making messages easily retrievable for listeners” (Slobin 1977:186) or as a tendency to express “each invariant, separately intuited element of meaning by at least one phonologically separate, invariant stress-bearing form” (Naro 1978:340). In SLA, this hypothesis competes with the traditional transfer hypothesis. That is: transparency predicts that L2 surface structures will tend to mark each significant underlying meaning element explicitly, irrespective of the facts of L1, whereas transfer predicts that L2 surface structures will directly reflect L1 structures. It must be noted that the two theories are not incompatible: it may well be that some transfer takes place but that the selection of the features transferred from L1 to L2 is determined by ST. In fact, it is commonly held that from most second language learners both ST and transfer play a role.

In the same way, the transparency hypothesis competes with the traditional substrate hypothesis, but not on a basis of incompatibility. For it may well be that the BL influences the creole in certain ways (in very many cases the evidence for substrate influence is overwhelming and highly convincing). But the selection made of the BL features that are carried over into the creole
may well be determined by the principle of ST, in that those features that have a high degree of ST are more likely to persist in the creole than low ST features. The two hypotheses get into conflict only when they claim exclusive rights.

As for the notion of ST, the important questions that remain to be answered include: what is meant by the concept of ‘semantic element’? How is the association between surface units and underlying meaning elements to be represented? And, are there any theories of semantic and grammatical analysis which allow the required one-to-one mapping between semantic structures and surface forms?

In the light of what we know today about semantic structures and semantic elements, it appears to be entirely unreasonable to think in terms of a condition of one-to-one mapping. Semantic structures are inevitably much richer than linguistic surface structures, mainly because they must be fully explicit and fully unambiguous. This particular condition is totally alien to the very nature of human language, in fact, if it were fulfilled, one would speak in some language of predicate calculus with totally unambiguous lexical items filling the places of the real variables.1) Such a language would not be a natural human language. The ‘definitions’ of ST quoted above would thus seem to be vague, unrealistic, and, finally, strictly theory-dependent in so far as the notion of “semantic element in a sentence” cannot be described without some specific theory of semantic analysis.

We feel that a more constrained definition is possible. But we recognise at the same time that any attempt at defining the or a notion of ST will of necessity be theory-dependent: after all ST involves a relation between semantic structure and linguistic surface structure, and thus presupposes a theory of semantic structure as well as a theory of surface structure. Now the latter is perhaps problematic in subtle and strictly technical ways, but the former is problematic in a much more profound and general sense. There is, as yet, no generally accepted theory of semantic structure. Agreement in this area does not go beyond the almost trivial condition that the language of semantics must contain the formal means for a logical calculus on analyses in terms of predicates and their arguments, plus quantifiers and logical connectives. Moreover, ordinary linguists are not usually in possession of the expertise developed by semanticists in the area of semantic analysis and semantic structure, due to certain historical accidents, in particular the dominant role of behaviourism in the human sciences between, roughly, 1930 and 1960. The semantic expertise in question is the property, generally speaking, of formal
semanticists, who are emphatically not linguists. (Their background is in logic and philosophy, and they have, generally speaking, no idea of the real complexities of natural language.) The in fact rather awkward attempts mentioned earlier at defining ST must be seen in the light of this historical perspective.

If ST is to be given a satisfactory definition at all, it must be in the context of a theory that makes explicit the relation between surface structures, on the one hand, and whatever seem to be reasonable semantic analyses, on the other. Any such theory will thus give an account of the relation between surface structures and what are selected as semantic structures. Or, in other words, any theory of ST will have to be formulated in terms of a grammar that defines the mapping relations between surface structures and semantic analyses.

In recent times various theories or paradigms have been developed to define this relation between surface sentences and semantic analyses. We shall, on the whole, remain as uncommitted as possible to particular theories. (But positions must inevitably be taken when specific examples are worked out and presented.) It should be observed that the vexed question of autonomous versus semantic syntax (Seuren 1972) need not plague us here, since it is immaterial to the point at issue whether there is or is not a separate compartment of grammar that is 'autonomous' with regard to the rule system relating semantic analyses and surface structures.

Intuitively speaking, ST can be seen as a property of surface structures enabling listeners to carry out semantic interpretation with the least possible machinery and with the least possible requirements regarding language learning. Surface structures or, rather, sets of surface structures forming a language, are higher on the scale of ST as their semantic interpretation requires less effort, both in the actual on-line processing and in terms of mastery of the system. If we wish to capture this intuitive notion in terms of mapping rules relating semantic and surface structures, we can say that a maximization of ST involves three strategies for grammars: (1) maximal uniformity of treatment of semantic categories, (2) minimal reliance on rules or rule types that are highly language-particular, and (3) minimal processing. Or, to put it briefly, uniformity, universality, and simplicity.

We now put forward the idea that creole languages are linguistically characterized by a tendency to maximize ST. Whatever has been or will be found to be typical of creole languages must at least be compatible with this condition, and much of what appears to be typical of creole languages will turn out to be epiphenomenal upon basic universals and their interaction with
situations where learning and processing efforts must be kept at a minimal level. It is intuitively easy to understand that creole languages should be characterized this way. The difficulty consists in making the notion of ST workable and giving it a functional place in existing linguistic theory. This paper is a first attempt at doing so against the background of modern grammatical and semantic theory. We shall comment on each of the three strategies distinguished above separately.

**Uniformity.** Given this strategy one will expect few arbitrary grammatical distinctions, as with grammatical gender or conjugational idiosyncrasies, or with derivational processes in morphology. Moreover, one will expect a uniform strategy for arranging verbs and their arguments (subject, object, indirect object). Thus, rules that bring about variations in the order of subject (S), verb (V), object (O), and, though less crucially, indirect object (IO), will be untypical of creole languages. For example, the rules responsible in German and Dutch for SOV order in subordinate clauses, as against SVO order in main clauses, are typically non-creole rules. Also, the rule of Predicate Raising, found in very many languages but particularly rampant in German and Dutch, is a typically non-creole rule, since it results after repeated cyclic application, in serial arrangements whereby the nominal arguments of the embedded clauses in question are strung together serially at one end of the clause, whereas the verbs in question are strung together serially at the other end.

In French, which likewise has Predicate Raising, though in a much more restricted form than Dutch and German, the set of verbs that take Predicate Raising is small compared with Dutch and German (five or six as against over fifty). Moreover, there appears to be a constraint in French against constituent crowding: the stringing together of nominal arguments from embedded clauses of varying depth, as found in German and Dutch, is not possible in French. French thus occupies a higher place on the ST scale yet it is far from optimal since the subject of the lower clause always ends up after the verb (as object, when the lower clause is intransitive, or as indirect object, when the lower clause is transitive). For the otherwise strong SVO-language that French is, this is an infringement of the uniformity principle. Now take Mauritian Creole, which also has the rule of Predicate Raising, obviously taken from French. But there, although the set of verbs that take PR is much larger than in French, the conditions of application of this rule are such that any violation of SVO-order is systematically avoided (cf. Seuren, in prep.). The significance of this is illustrated by the fact that in an ordinary synchronic syntactic description of Mauritian Creole the restrictions on Predicate Raising will
form a odd-looking and improbable collection, until it is seen that they con-
spire to maintain pure SVO-order in all cases. In this creole language it never
occurs that a semantic subject follows its verb due to Predicate Raising.

*Universality.* A maximization of semantic transparency in the sentences
of a language will mean that maximal use is made in the grammar of those
processes and rule systems that are least language-specific, so that the uni-
versals which leave the least room for language-particular variation are exploited
most. This renders morphology essentially alien to creole languages, since
whatever universals enable the growth of a morphological system in a lan-
guage leave ample room for a multitude of often haphazard variations. One
might observe, in this connection, that languages with a complex morphology
are generally hard to master for non-native learners, even when the mor-
phological system is semantically highly regular ('uniform'), as is the case
with, for example, Turkish.

Another consequence of the universality condition is that so-called 'sec-
ondary constructions' will seldom be found in creole languages. A secondary
construction (cf. Stassen 1985) is grammaticalized only in older or more
advanced languages, but its semantic content is expressed by normal creative
and ad hoc means in younger or less advanced languages. A striking example
of a secondary construction is the comparative, which, in many beginning lan-
guages, is not grammaticalized at all but expressed by ad hoc means, such as
“A is big from B”, “A is big to B”, “A is big, B is not big”, etc. Serializing lan-
guages typically use serial verbs for the expression of comparatives: “A is big
exceed B”. Nominalizing languages typically use constructions of the type “A
exceed B in bigness”. In the grammars of languages that have no gram-
maticalized comparative, there is no chapter called “the comparative”, since
whatever is to be expressed is expressed by means of independently available
grammatical constructions. It is not known how comprehensive the set of sec-
ondary constructions is in the languages of the world. But the prediction is
that creole languages will have an extremely small share of them.

*Simplicity.* This criterion implies that the amount of processing needed
to get from semantic analyses to surface structures, and vice versa, is kept to a
minimum. This minimum appears to be solid, in the sense that it seems
impossible for any language to reduce the amount of processing to zero. As
we have said earlier, this would mean speaking in some predicate calculus
language. All semantic theories agree that something like predicate calculus,
with bound variables and the rest, must determine the structure of semantic
analyses. Yet no language has surface sentences maintaining any such struc-
The rationale for this in itself striking fact seems to be the following. Predicate calculus structures expressing what sentences mean are characterized by multiple embedding of sentential structures and thus depend essentially on their tree structure (bracketing structure) for their proper interpretation. Surface sentences, however, do not, carry their tree or bracketing structure with them in acoustic or even written transmission. Rapid acoustic transmission requires a phasing out of the role of tree structure, which must be replaced by different, acoustically perceptible means, in particular word order and intonation. This being so, a minimal amount of transformational processing is anyhow required, but it still remains to be seen what this minimum actually amounts to (hardly any research effort has been devoted to this question). However, as far as the major rules are concerned for the lowering of logical operators (negation, conjunction, disjunction), of quantifiers (so that bound variables are eliminated or replaced by bound variable pronouns; cf. Seuren 1985, ch. 4), of sentence adverbials, tenses, and modalities (if they are lowered and not treated differently) it would seem that these forms of lowering are obligatory in all languages. But other cyclic rule systems, such as Subject Raising, or Predicate Raising, are dispensable.

Postcyclic rules resulting in the movement of major or minor constituents within the sentence, or in deletions (such as Gapping) are against the spirit of the simplicity condition and are thus expected to occur only rarely in creole languages. An exception must be made for postcyclic rules reinforcing semantic salience. For example, the question-forming property of specific question words (WH-words) must be considered semantically salient, and it is therefore to be expected that this function is somehow highlighted. A common strategy for highlighting this function is WH-fronting, found in very many languages all over the world, and also in creole languages. (In general, the principles of uniformity, universality and simplicity can be counteracted or overridden in cases where semantic salience is to be given its due to the perceptible sentences form. In such cases iconic expressions (cf. Haiman 1980, 1983) will be preferred, to minimize the ‘damage’ done. Reduplication, for example, which is an eminently iconic expressive tool, will be used for the expression of plurality where this is necessary, or for the expression of iterative, hesitative, or continuative aspect).

It is generally to be expected that the surface form of creole sentences will differ relatively little from their shallow level representation, i.e., the structure resulting from the application of the cyclic rules and before any
postcyclic rules apply. Moreover, the elements lowered will reflect in their left-to-right order the semantic scope that is expressed by means of tree structure in semantic analysis: there will be a direct, one-to-one relation between scope and left-to-right order. This, again, leads to a ban on morphology, since, as is well known, morphological processing of lowered elements loads to a massive violation of this scope-order correspondence, and thus requires a great deal more cognitive processing than is needed for sentences with a regular scope-order correspondence. In this light it is not surprising to find that in many creole languages verbal tenses, modalities and aspects are expressed by means of preverbal particles, and not by morphological means, as in the majority of more advanced languages. The occurrence of such particles is a direct reflex of their semantic scope after lowering.

Semantic transparency, in an overall sense, limits learning and computing efforts to a minimum. The price to be paid for this is more effort: the sentences will be more circumstantial and at the same time the total repertoire for semantic expression is more restricted. More advanced languages require more learning effort and more computation for the interpretation of their sentences. But they can do with less verbal material, and have richer expressive means. This, it would seem, is part of the functional picture arising in this context. But it must be added that other factors play a role as well in the trade-off between instant labour and 'investments' in computing apparatus. Languages seem to allow themselves certain 'luxuries' in the form of idiosyncratic exceptions, or highly specialized lexical items or their uses. Such 'luxuries' seem to be provoked by social differences within the speech community, where speakers of higher rank have an interest in making their speech hard to imitate. But in a situation of incipient creolization, such luxuries will hardly be found.

The perspective of semantic transparency as a determining factor in creole genesis opens up a massive research programme, integrating creole studies into the general theory of grammar and semantic expression. We feel that such research is long overdue and will be highly profitable to both the theory of creole languages and the general theory of language which has not so far been seen to take an active interest in questions of linguistic variation over time, or, generally, in questions of what can be considered the 'life' of language systems in communities of living speakers. If the notion of semantic transparency can indeed be made workable, precise and operational, this may help the currently dominant general theories of language, which are too often characterized by sterile idealization, to combine the theoretical power and sophistication which they undoubtedly possess with a greater sensitivity
to the enormously complex empirical reality of natural language. An important step would thus be taken to approximate the ultimate goal of a formal and explanatory theory of not just an idealized abstraction of language but of something much closer to full and ordinary reality.

NOTE

1) By "real variables" we mean the variables that occur in logical calculi and range over non-logical constants. Thus, a Predicate Calculus may have the variables $K, L, M$ for predicates, $a, b, c$ for individuals; or a propositional calculus can use the variables $p, q, r$ ranging over actual propositions. Real variables are opposed to technical variables, which have an internal technical function within the calculus. In practice, technical variables are either free or bound variables; they are usually represented by the letters $x, y, z$.

REFERENCES

Seuren, P. in prep. Notes on the history and the syntax of Mauritian Creole.
Slobin, D. 1977. Language change in childhood and in history, in J.

P. Seuren & H. Wekker
Filosofisch Instituut KUN/Instituut Engels-Amerikaans KUN
Erasmusplein 1
6525 HT NIJMEGEN
Netherlands