Review Article


Abstract
This review article consists of three parts: a presentation of McWhorter’s book, an evaluation of the book, and a proposal for a systematic metric for one version of the notions of simplicity and complexity in languages. McWhorter’s basic thesis, that creoles have simpler grammars than their lexifier languages and simpler grammars than languages transmitted through acquisition in childhood, must begin with rigorous definitions of the terms ‘simplicity’ and ‘complexity’. The paper proposes a way to measure simplicity and complexity. Whether the notions of simplicity or complexity have a heuristic value remains an open question.

Keywords
Creoles; language simplicity and complexity; language contact; language change

The aim of this review article is to present McWhorter’s book; pose a few questions; and to provide a different metric for the title theme of this book, the notion of simplicity and complexity. This review will not attempt to referee disputes between McWhorter and opposing creolists, of which there appear to be quite a few and which are not always of interest to general linguists, and it will not match McWhorter’s elegant and engaging style of writing. This review does, however, address some of the issues of interest to general linguists. Hopefully, it will also provide some precise formulations for notions that McWhorter invokes but does not thoroughly explain, thus providing support for some of his ideas.¹

¹ I would like to thank Erin Shay for her careful reading of this paper and for her numerous comments re: organization, substantive issues, and the style of the text. Her advice, as always, is invaluable.
1. Basic questions of the book

While much of the book is dedicated to rebuttals of arguments against McWhorter’s earlier writings, the book deals with a number of issues worth discussing.

It appears that the following are research questions that McWhorter is trying to answer with respect to creoles:

(1) Are there grounds for proposing a new typology of languages that distinguishes between languages that are acquired in childhood and those that are acquired in adulthood as a second language? The latter is McWhorter’s definition of creoles, which replaces the historical definitions used so far. English, in McWhorter’s discussion, belongs to the group of languages that once were acquired as a second language.

(2) What are the types of changes that languages undergo when they are acquired in adulthood?

(3) What are the reasons that certain languages are typologically different from their sisters?

In McWhorter’s approach, the answers to all of these questions hinge on one and only one process: When a language is acquired in adulthood it undergoes simplification, its inflectional morphology is lost, and it does not display irregularities. Consequently there are two types of languages: ‘old’, i.e. languages undergoing the usual changes that languages undergo when acquired in childhood, and ‘new’, i.e. languages acquired in adulthood. In McWhorter’s view, languages acquired in adulthood are simpler than their sister languages acquired in childhood.

2. Presentation of the book

The book is composed of three sections and an introduction. The introduction consists of the basic concepts of McWhorter’s approach to the issues of creole languages. Sections I and II consist of rebuttals of numerous criticisms of McWhorter’s claim that creoles have simpler grammars than the languages from which they are derived and the claim of creole exceptionalism, i.e. the existence of creoles as a linguistic type, regardless of the history of their development. Section III applies this approach to Riau Indonesian, the languages of Flores island in Indonesia, and English. Except for the introductions, the book consists of the development of papers published in a variety of journals during the years 2008-2012. Though the book addresses issues that evidently
raised much discussion among creolists, it is not without interest for general linguists. The recurring issues of the book, all of which are described in the introduction, are:

(1) The notion of complexity and simplicity;
(2) Creole exceptionalism; and
(3) The defining characteristics of creoles.

The fundamental issue of the book is the notion of complexity. In McWhorter’s approach, complexity is characterized by the following:

**Overspecification.** This apparently fundamental characteristic is not defined. Instead it is illustrated by examples involving obligatory overt marking of such semantic distinctions as ‘inalienable possession’, ‘the ‘fourth person’ obviative, evidentiality, and definiteness’ (p. 2). While one may intuitively understand what McWhorter means by overspecification, given the importance of this characteristic for the whole work, one would expect a rigorous definition rather than merely a few examples. In the absence of such a definition, overspecification cannot be used as a criterion of complexity. The choice of the term overspecification is particularly unfortunate, as it is an inherently comparative term. It is not clear what the standard is for classifying some features of a language as overspecification.

**Structural elaboration.** Here I must cite McWhorter’s formulation, so as not distort his intention: ‘An aspect of one grammar may differ from that aspect in another’s in terms of the number of rules (in phonology and syntax) required to generate grammatical forms; examples include morphophonemic processes, concord, and heterogeneous word order’. This description of ‘structural elaboration’ assumes the existence of underlying forms for morphemes, an assumption shared by many linguists. This description also assumes some kind of underlying syntactic structures, an assumption shared by considerably fewer linguists. Moreover, the description relies on unspecified assumptions about specific grammatical terms, as reflected in the use of the terms ‘concord’ and the undefined ‘heterogeneous word order’.

**Irregularity:** ‘grammars differ in the degree to which they exhibit irregularity and suppletion’ (p.2), with the irregularity being characteristics of old languages.

For this reader, this description of complexity appears to be ad hoc rather than derived from a particular language theory.

The rather vague characterization of complexity is followed by a bold statement that languages that ‘are significantly low on this kind of complexity compared to their sisters owe this state to second-language acquisition in the past’, and a follow-up claim that ‘languages with the least complexity of all the world’s languages are creoles’ (p.3).
Section I is dedicated to the issue of creole exceptionalism. It consists of an introduction and three chapters, each of which is a rebuttal of arguments made by various linguists against the claim that creole languages are typologically, not historically, different from other languages. In chapter 1, McWhorter revises his notion of Creole Prototype that he proposed in his earlier writings, and acknowledges that individual creole languages may differ from the prototype with respect to one or more of the characteristics listed above, but never with respect to all three characteristics.

With respect to inflectional morphology, McWhorter adopts Booij’s distinction between two types of inflection, one coding such categories as tense, mood, and aspect, called ‘inherent inflection’, and the other coding case and concord, called ‘contextual inflection’. According to McWhorter, the creole languages may have inherent inflection but not contextual inflection. For McWhorter, inflection, following Kihm 2003, includes the traditional marking through the changes on the word as well as free morphemes. The revised Creole Prototype, as presented in McWhorter’s earlier writings, now includes the following:

Morphology: Little or no inflectional affixation, no contextual morphology.
Phonology: Little or no distinction of monosyllabic lexical items through tone, and no typologically unusual proliferation of vowels.
Semantics: Little or no noncompositional combination of nonreduplicative derivational morphemes with roots.

In McWhorter’s view, all of these characteristics are the result of ‘young’ (creole) languages emerging from structurally reduced pidgins. (p. 60-61)

Chapter 2 is a rebuttal of arguments against McWhorter’s reactions to the treatment of creoles in various papers in Plag 2003. The specific aim of this chapter is to demonstrate, against his critics, that in all areas of grammar creole languages are less complex than their lexifier languages and, in general, less complex than old languages. The method of rebuttal consists in taking some paradigm or grammatical fragment from an old language and comparing it with a paradigm coding the same function in a creole. This is done with respect to the future tense in Russian and Saramaccan (pp. 66-67), with the predictable result that Russian makes more distinctions with respect to person and number and that the form of the morphemes involved is less predictable in Russian than in Saramaccan. Complexity in phonology is illustrated by contrasting phonological rules in Kabardian, an old language and Haitian, a young language, with Kabardian having more rules for deriving the phonetic realization from the underlying structure (pp. 74-75). Complexity in morphology is demonstrated, by contrasting Tok Pisin marker -pela with the partitive marker in Finnish, with the latter having more irregularity than the morphology of Tok Pisin; and the comparison of the plural marking in Nuba
Arabic with the plural marking in classical and contemporary Arabic dialects. The main conclusion is that older languages have more morphemes marking more categories than young languages. However, the chapter does not provide systematic evidence for this conclusion.

Chapter 3 is a rebuttal of arguments against McWhorter’s approach in papers in Ansaldo et al. (eds.) 2007. The main issue again is the problem of creole synchronic exceptionalism resulting from the fact that creoles are historically newer than their lexifier languages.

Section II is dedicated to the issue of complexity and consists of three previously published papers on Saramaccan. The main thrust of the section is to convince (unfriendly) creolists, and possibly other readers, that McWhorter does not claim that creole languages have no grammar. The three papers describe the grammaticalization of various functions, including new information markers; serial verb constructions with equivalents of the verbs ‘throw’ and ‘pull’; and the complexity of the Saramaccan copula. The discussion of the latter gives McWhorter an opportunity to address the issues of the metric of complexity. He illustrates the notion of ‘overspecification’ on the basis of Kwaio (an old language, from Keesing 1985), which has three persons and four numbers (singular, dual, paucal, and plural) in its pronominal system. From the way the forms are presented, it appears that there is a distinction between inclusive and exclusive forms in the first-person singular, dual, paucal, but not in the first-person plural. Given the differences in the forms of the pronouns, this distribution does not appear to be the result of a typographical error, but the functional difference between the first-person singular inclusive and the first-person singular exclusive pronoun is an intriguing question.

The illustration of structural elaboration is based on a discussion of Welsh consonant mutation, which, according to McWhorter, requires a large number of rules to generate the attested grammatical forms. The importance of the copula in Saramaccan is that it illustrates that complexity can indeed be generated through grammaticalization processes within a creole language.

Section III discusses exceptional language change in Riau Indonesian, in languages of the Flores island, and in English. With respect to Riau Indonesian, McWhorter claims that this language has characteristics that indicate that it was originally acquired as a second language. Some languages of Flores differ from other Austronesian languages in that they are affixless and completely isolating. Here the claim is again that they have become affixless as a result having once been acquired as second languages.

The last chapter of the book deals with the Celtic influence on English, and more specifically with the sources of the auxiliary ‘do’ in English. It is a lengthy and carefully written chapter in which McWhorter reviews the
previous scholarship on Celtic influence and asserts, in agreement with van der Auwera and Genee 2002, that the auxiliary ‘do’ emerged as a result of contact between English and Cornish. The argumentation for this hypothesis is that although the equivalents of ‘do’ have been grammaticalized in various languages as auxiliaries, the functions of this auxiliary in English are similar to those in Cornish. The similarity is not complete, as Cornish uses the equivalents of the auxiliary ‘do’ in the affirmative, while English uses only in negative and inverted clauses. The similarity between Cornish and English poses the question of directionality, to which McWhorter devotes considerable amount of space. The problem is compounded by the fact that Cornish attestations of the auxiliary ‘do’ are later than the first attestation of auxiliary ‘do’ in English. Nevertheless, McWhorter proposes that the auxiliary ‘do’ in English is a product of language contact with Cornish. In McWhorter’s view, English bears many characteristics of a language acquired as a second language by Celtic speakers, in the process of which it shed much of it inflectional characteristics.

3. Evaluation

This is a carefully prepared volume. Even though much of the material has been previously published, the volume and each section in the volume has a new introduction that sets the background for things to come.

McWhorter’s fundamental thesis, that creoles do not reproduce the entire grammar of their lexifier languages and that this follows from the fact that creoles developed from pidgins, which themselves had reduced grammatical systems in comparison with the languages from which they developed, is indeed non-controversial. As anybody who has had to learn to speak a second or third or fourth language in adulthood will confirm that such an activity results in considerable reduction of the new grammatical system. But what gets reduced is an open question. It appears that the first elements to be reduced are categories that do not exist in the speaker’s native language. For example, it took this writer a considerable amount of time to properly incorporate the perfect aspect and the so-called progressive aspect into his English usage, as neither of these aspects exists in the grammatical system of Polish. There may be also other elements that are reduced. What they are, and why they are reduced, remains an open question.

Contrary to McWhorter’s claim, language acquisition in adulthood is not the only reason for reduction of the affixal system and for drastic changes in the grammatical system. Phonological changes, such as shift of stress, may
cause the weakening of final syllables. This may lead to the loss of final vowels or consonants. If those vowels or consonants happen to be grammatical markers, the reduction of the syllable may lead to the reduction of much affixal morphology. This is what happened to French as it evolved from Latin.

The issue of the Creole Prototype may be approached in two ways. One is by comparing creole languages to their lexifier languages. This is what McWhorter does. The other is to compare all languages that are generally agreed to be derived from former pidgins, see what they have in common, and then compare those features with languages that are generally agreed not to be derived from pidgins (‘old’ languages, in McWhorter’s terminology). McWhorter does this with respect to just three characteristics, viz. the absence of affixal inflectional morphology and the absence of inflectional morphology coding grammatical relations; the absence of non-reduplicational derivational morphology; and the absence of tone on monosyllabic morphemes and the absence of proliferation of vowels. Creole languages are also said to be characterized by the absence of irregular morphology, as represented by the English verb ‘understand’. Indeed, its morphology is obscure.²

Why these three characteristics, and not others, were chosen as the main criteria appears to be the outcome of some empirical investigation and a reaction to previous criticism. In what follows, I propose a way of conducting a systematic study of typological characteristics of languages. If creoles emerge as a group from such a comparison, McWhorter’s hypothesis will be vindicated.

The investigation of complexity, an issue that is the subject of a number of contemporary studies, hinges crucially on a theoretical definition of this term with respect to language structure, which in itself is a complex entity. When McWhorter claims that creole languages are less complex than their lexifier languages or less complex than old languages, the absence of a clear theoretical definition understandably provokes many creolists. Who would be happy to acknowledge that they work on less complex languages than other people?

As McWhorter states, language acquisition by adults, which is almost always imperfect, can also result in the loss of inflectional morphology. The loss of inflectional morphology may have several sources, not just one, as claimed by McWhorter. To extend McWhorter’s metaphor, ‘languages

² The present writer was kicked out of a beginning English class when he used the verb ‘understand’ in the sense of ‘do not understand’, thinking that the verb consisted of the negative marker ‘un’ and some stem presumably carrying the affirmative meaning corresponding to ‘comprehend’.
undress’ because the cloth gets worn out in all kinds of places, but most often at the edges.

4. Toward a systematic metric for complexity

4.1. The notion of complexity

The notion of ‘complexity’ in any system has to be defined for each system and cannot be used aprioristically. One component of complexity in language is the number of coding means the language has. Another component is the number of functions coded in the language. The two numbers are not the same, because coding means can be combined in different ways to code various functions; the number of functions is therefore greater than the number of coding means. Moreover, some coding means have the potential to produce a larger number of constructions than do other coding means. Thus, tone can drastically increase the number of functions when it is combined with segmental morphemes. With respect to monosyllabic morphemes it can double the number of potential distinctions, and of course the number of potential distinctions increases with longer words. Given that a single function can be coded by a variety of constructions, the third component of complexity is the number of constraints that affect the realization of a given function. Such constraints include properties of lexical items chosen for the realization of a given function and the interaction of a given construction with other constructions used in the same utterance. Below is a brief outline of coding means and functions they encode, and examples of the interaction of the coding means in the realization of a given function.

4.2. The coding means

The actual coding means must be determined for each language. The coding means frequently available in languages include:

Phonological alternations
Prosody
Lexical categories and subcategories
Derivational morphology
Inflectional morphology on all lexical categories: Nouns (includes not only case marking but also the marking of states in Semitic languages and a completely different category of marking, also called ‘state’, in Berber); verbs
includes marking of arguments and adjuncts as well as verbal extensions),
adjectives, numerals, adpositions, complementizers, etc.
Linear orders (as outlined in Frajzyngier 2011)
Serial verb constructions
Grammaticalized lexical items, such as body-part terms coding spatial
configuration
Adpositions
Conjunctions, complementizers, and subordinators (see Frajzyngier 1996)
Repetition of individual lexical items or even phrases (cf. ‘progressive aspect’
in Gidar (Chadic), Frajzyngier 2008)

4.3. Functions

The functions of various coding means, like the coding means themselves,
have to be discovered for individual languages. Contrary to McWhorter’s
claim, overspecification or for that matter underspecification, does not exist as
a comparative concept. The way one can compare the functions encoded in
individual languages is by comparing the number of functional domains that
the language encodes and the internal structure of those functional domains.
Thus, some languages have the domain of tense but not aspect, others have
aspect but not tense, and still others have both the domain of aspect and the
domain of tense. From the point of view of functions, it doesn’t matter how
these functions are coded, whether through inflection on verbs; on nouns, as
is the case in some Australian languages; through serial verb constructions; or
through auxiliaries. Even if two or more languages have the same functional
domains, the internal structure of these domains may differ significantly.
Thus, in the domain of equational predication of the type ‘X is Y’, some lan-
guages make no distinction between permanent and temporary states, while
other languages, e.g. some varieties of English, do make this distinction.

Once one has a complete list of the formal means and functional domains
in a large number of languages (no such list exists for any language, as of yet),
it will be possible to develop some types of language hierarchies, based on
which languages have more coding means and which have fewer coding
means, which languages have more functions encoded in the grammatical sys-
tem and which have fewer functions. It will then be possible to decide whether
there is a correlation between the number of coding means and the number of
functions coded.

But even if one finds that some languages code more functions in their
grammatical system than other languages, this does not necessarily mean that
these languages are more complex in speech processing. The third factor that
may affect the perception of complexity is the interaction of various coding means in the coding of a function.

4.4. The interaction of coding means in the coding of a function

Although a one-to-one correlation between a coding means and a function is possible, many functions are coded by a combination of a variety of coding means. The interaction may depend on the properties of the lexical items deployed in the coding of a given function and/or other functions coded in the utterance. An example of the interaction of lexical properties with grammatical means is provided by the English adverbial function. When a lexical item is an inherent adverb or a morphologically derived adverb, the adverbial function is coded only by the lexical item:

Coincidentally, some say that being Mitt is easy. (COCA)

If the lexical item is not an inherent or transparently derived adverb, the adverbial function must be marked by a preposition in addition to the clause-initial position:

By coincidence, Whittaker (1957) published his first article on kingdoms just a few months before the launch of Sputnik 1, (COCA)

Another example of the interaction of coding means is provided by the coding of interrogative modality in a language with rich inflectional coding on nouns and verbs. In such a language, linear order codes a variety of functions in the domain of information structure, and therefore is not available for the coding of interrogative modality. Hence, the interrogative modality is coded by particles (in addition to prosodic means), as is the case in Polish:

Czy zbliża się Rzeczpospolita?


An example of the interaction of various coding means motivated by the presence of different functions in the utterance is provided by the coding of interrogative modality in embedded clauses in English. Grammatical relations in the embedded clause, like those in the matrix clause, are marked by the linear order SVO. Subject Auxiliary/Verb inversion is the means of coding interrogative modality in the matrix clause. If the same means were applied in the embedded clause, the language would lose the means of distinguishing between two consecutive sentences in discourse and the embedding of one clause within another, which forces the listener to interpret the two clauses
together. This explains the obligatory insertion of the marker ‘whether’ to code the either/or\(^3\) questions, in the embedded clause (all examples from COCA):

Some researchers have begun questioning whether an automatic system is the right choice for this task.

Other ethical questions likely to arise include whether it’s “right” for people to fall in love with (or even marry) their robot partners.

For months listeners debated whether the husband could have done it.

How would he decide whether to bomb Iran or let Israel bomb Iran?

Once one has developed a complete list of formal coding means existing in a language, a complete list of the functions coded, and a complete list of the potential interactions of the formal means, one can then compare languages with respect to these three parameters and then label the languages (albeit arbitrarily) that have a larger number of coding means, functions, and interactions as more complex and languages that have a smaller number of these characteristics as less complex. Unfortunately, such labeling has no heuristic value. However, trying to explain why some languages have more or fewer formal means and more or fewer functions is a worthwhile enterprise.

5. Conclusions

McWhorter’s book is thought-provoking and well written. Although it is mainly a polemic with other creolists, it is of interest to all students of language contact and to general linguists.

Zygmunt Frajzyngier
University of Colorado
Email: Zygmunt.Frajzyngier@Colorado.Edu

References


\(^3\) I am grateful to Erin Shay for pointing out the either/or function of ‘whether’ in the embedded clauses.
COCA THE CORPUS OF CONTEMPORARY AMERICAN ENGLISH http://corpus.byu.edu/coca/


