The grammaticalization of lexical items in pidgin/creole genesis

ANTHONY J. NARO
Universidade Federal do Rio de Janeiro

There are many possible scenarios for the beginnings of pidgin/creole genesis, but they all have in common a need to communicate in a situation in which there is no common language among the people involved. The situation can be momentaneous or sporadic, such as contact between Norwegian and Russian fishermen in the North Sea in the last century, or it can be drawn out and continuous, such as contact in Portugal between the Portuguese and African slaves brought directly from the African coast by the first exploratory expeditions beginning in the mid 15th century.

In most of these situations one of the groups involved is socially and economically dominant. In the initial period of European overseas expansion, as well as during the occupation of the Americas and the colonial era, for example, the Europeans were the masters and overlords, imposing their will on everyone else. When we consider the linguistic state of affairs that took hold overseas, we see that the Europeans did manage to impose their own vocabulary, or set of lexical items. In fact, the lexicons of most modern day pidgin/creole languages, from the Caribbean to the Indian Ocean, are composed of more than 90% of items directly traceable to European languages. For this reason, the language spoken in Haiti, for example, is called Creole French and is sometimes classified as a French dialect, simplified French, bastardized French, or with other similar labels.

It is a well known fact that the syntax and semantics of the pidgin and creole languages do not bear the same degree of similarity to the European languages that is so apparent in the case of the lexicon. But the lexicon too, of necessity, underwent drastic changes. For one thing, in the first stages of contact, when neither side knew anything of the other's language, the lexicon was vastly reduced in terms of the number of items used. Even as contact progressed, the range of social contexts in which the contact language was used were usually quite limited. For example, although China Pidgin English existed for many decades, its use was limited mainly to trade, and its total vocabulary probably never exceeded seven to eight hundred items.

In most areas, the language that resulted from contact was appropriated by those from the socially inferior side. This was necessary because, for diverse reasons, these people had no language in common among themselves. In the Americas, for example, slaves were brought from vast areas of the African

continent, and there is evidence that a deliberate effort was made to mix people of different languages precisely in order to hinder their ability to communicate among themselves and thus lessen chances of rebellion. In other cases, such as in China and New Guinea, the economic activity created by Europeans attracted people from such far-reaching areas that the new arrivals had no common language that they could use to communicate among themselves. Under these conditions, the contact systems based on European lexicons were taken over by the other side and came to be used in all of the contexts typical of daily life. Their environment of use was no longer limited to whatever happened to interest the dominant side in the contact of superior with inferior.

When the social situations in which the contact language was used underwent expansion, the lexicon had to expand as well. However, the dominant Europeans were too far socially removed to be the source of more lexical items, and, as we have seen, those from the inferior side had no common language from which to bring in new lexical items. The result of this is that the lexical items that existed had to be stretched to meet the communicative needs of the population. This occurred in three ways: (1) the creation of standardized phrases to take the place of non-existant lexical items, (2) the attribution of new meanings to the relatively small number of lexical items in use, (3) the attribution of new syntactic and morphological roles to these items.

The first techniques, the introduction of new fixed expressions, resulted in phrases that seem picturesque from the European point of view: *grass belong head* for 'hair', or *ass belong world* or *big fellow* for 'God' in New Guinea, for example. The second method, enlarging of semantic fields, resulted in certain lexical items gaining far wider semantic ranges than in the European languages. A typical case is the preposition *na*, sometimes thought to be derived from standard Portuguese *na* 'in the (fem.)', used in many pidgin/creole systems of diverse lexical bases with a extensive range of meanings including 'to', 'from', 'in', 'into', 'out of', etc. Another example is found in New Guinea pidgin *two-fellow moon he-die* for 'two days passed', in which the lexical item *die* (as well as *moon*) has acquired an additional sense not present in the standard.

It is to the third method of lexical expansion, change of morphosyntactic role, that we will turn our attention in this paper. Under this category one usually has in mind examples such as change of the preposition *behind* to an adverb meaning 'later' in New Guinea pidgin. An example is:

Behind by-'n-by he come-back again

for 'later it will come back again'. Another case of the same type is the use of a noun as a verb in

Boatswain gammon me

for 'the boatswain lied to me' (literally 'boatswain game me').

Here we will examine a change of category of a more radical type, where a lexical item looses its semantic content and becomes a functional item within the grammar. This type of change is called 'grammaticalization'. A typical case of this sort can be found in the history of the future with *gonna (going to)* in standard colloquial English. Originally a construction such as

I am going to read my paper

had a purposive meaning involving actual physical dislocation in space along the lines of 'I am traveling (from Rio de Janeiro to Granada) in order to read my paper'. From this meaning the future developed: if I intend to do something, that thing hasn't happened yet, i.e., it is in the future. The modern English *going to* construction now has both meanings, and they can sometimes be difficult to distinguish. The phonologically reduced auxiliary verb form *gonna*, however, can have only the grammaticalized future meaning, in which the lexical content of physical motion associated with the verb *to go* is no longer present. In fact, it is now possible to use both *going to* and *gonna* with clauses in which the semantic element of purpose or intent is virtually impossible and the future is the only reading possible. An example is:

I am going to (gonna) hear the chime when the alarm goes off.

The common Romance future (Portuguese: *cantarei*; Spanish *cantaré*, French *je chanterai*, etc.) also developed through a process of grammaticalization from the construction *cantare habeo*, with loss of the lexical content of possession in *habere* and the later extension to intransitive verbs. In this case the phonological reduction to the former lexical item, the verb *habere*, has proceeded further to a clitic stage, not yet attained in the case of colloquial English *gonna*.

One of the most intriguing aspects of pidgin/creole grammatical structure is the verb and its associated categories of number and tense/aspect. In general, lexical items such as nouns, verbs, adjectives, and so on tend to have only one invariant phonological form in pidgin/creole languages. In the case of verbs the most important consequence of invariance of form is the fact that the tense/aspect and number elements of meaning must be expressed by means of independent lexical items rather than through bound inflections. The verbal root itself does not express any particular tense or aspect. When used in a clause, it is up to the listener to interpret this aspect of meaning in accord with the context. In a Portuguese-based pidgin, for example, *sol nascer* might mean 'the sun comes up', 'the sun came up', 'the sun will come up', 'the sun is coming up', 'the sum was coming up', 'the sun will be coming up', and so on, according to the discourse context in which it occurs. In case the context is not clear in this regard, the

speaker might add an adverb such as *agora* 'now', *ontem* 'yesterday', *amanhã* 'tomorrow', an adverbial phrase such as *daqui a pouco* 'in a little while', or an adverbial particle such as *já* 'already' or *logo* 'right away'. This would clarify the intended meaning.

In the 1970's, New Guinea Tok Pisin was undergoing a period of expansion and stabilization as a language used in the full range of social situations of daily life. In its earlier stages, the future in Tok Pisin was indicated, optionally, by the use of any lexical element with a future meaning such as the adverbs *close-to* 'soon' or *behind* 'later', or the more general and vague adverbial expression *by-'n-by* 'in a while', usually in sentence initial position. Sankoff & Laberge (1982) studied the expansion process and found that *by*, a shortened form of *by-'n-by*, was beginning to take on properties typical of a grammatical future marker. The characteristics they list as differentiating the new grammaticalized future marker *by* from the previous lexical adverbial expression *by-'n-by* are: (1) reduction in phonological material, from the full bisyllabic form to only one syllable, frequently with a reduced shwa-like vowel; (2) reduction in stress, from the earlier primary word stress level to that typical of unstressed syllables in full lexical categories; (3) redundant use of *by* even when the future meaning is conveyed by lexical items such as adverbs; (4) tendency toward placement of *by* immediately before the verb, rather than in sentence initial position, especially with full noun phrase subjects. In this way the full lexical expression *by-'n-by* is converted to the reduced grammatical item *by*. In other words, grammar arises from the lexicon.

In Papiamento, as well in many Portuguese-based creoles (principally in Asia), the future or, in more technical terms, the irrealis mood, is indicated by means of the grammatical particle *lo*, derived from the lexical adverb *logo* 'right away'. Like *by* in New Guinea, *lo* is unstressed and is used redundantly even when the semantic content of futureness is already established by lexical items present in the clause or extra linguistic context. Although we do not have a fully documented history of Portuguese pidgin, it is possible to establish earlier usage by examining the imitation of the pidgin found in certain 16th century plays. One particularly interesting example, found in the anonymous play *Auto dos enanos*, is a scene in which two Portuguese peasants are trying to communicate with a Castilian noblewoman. No Africans or other non-Europeans are present in this particular scene, although there were many Africans in Portugal at the time. The Spanish lady is consistent in always using standard Castilian, but the Portuguese peasants are not very good at understanding her. At one point, she complains, saying to the peasants "You don't understand", to which one of them replies:

Nam, senhora.	No, madame.
Se vós falar como nós	If you speak like us,
logo nós entender vós	We will understand you
& responder *logo* ess' ora.	And reply immediately.

Notice that all of the verbs in the Portuguese peasant's speech are in the infinitive and the pronouns are all fully stressed non-clitic forms, in typical pidgin fashion. There are two occurrences of *logo* in the peasant's speech. The first is used as a general future marker, in a way quite parallel to *by-'n-by* in New Guinea pidgin. The second, with the original meaning of *logo* 'right away', is reinforced with the expression *ess' ora* 'right now' because the meaning of *logo* had already been weakened to the point where it no longer contained the lexical semantic element of immediateness. In the course of time, *logo* was phonologically reduced to just *lo* and became completely detached from the lexicon and integrated into the morpho-syntax, in a way parallel to the development of *by* from *by-'n-by*.

We see then, that the beginning of the pidginization process consists of a reduction and simplification on the linguistic levels of the lexicon and the grammar, parallel to the narrowing of the social contexts in which the pidgin is used to the interaction of lower and upper social strata. When the social situations of language use expand upon the appropriation of the language by a community of speakers for use in the full range of situations typical of daily life, the lexicon and grammar must expand as well, in order to adequately fulfill the new functionality. One of the ways in which this occurs is the transformation of full lexical items to morpho-syntactic items, a process known as grammaticalization. In effect, the lexicon suffers a minor loss in order to expand the grammar. The lexicon's loss is only temporary, however, as it too undergoes expansion through the processes outlined above. In the long run, a new full language, capable of fulfilling all of its social roles, is created from the limited resources of the pidgin.

REFERENCES

HOPPER, Paul J. & CLOSS TRAUGOTT, Elizabeth (1993) *Grammaticalization*. Cambridge: Cambridge University Press.
HOLM, John (1988) *Pidgins and creoles*. Volume I: Theory and Structures. Cambridge: Cambridge University Press.
NARO, Anthony J. (1978a) "Pidginization, creolization, and natural change". *Revista Brasileira de Lingüística* 5:123-138.
NARO, Anthony J. (1978b) "A study on the origins of pidginization". *Language* 54: 314-347
NARO, Anthony J. (1998) "O uso da concordância verbal no português SUBSTANDARD do Brasil: atualidade e origens", in *Substandard e Mudança no Português do Brasil*, ed. by Sybille Große & Klaus Zimmermann. Frankfurt am Main: TFM, 1998.
NARO, Anthony J. & VOTRE, Sebastião J. (1998a) "Mecanismos funcionais do uso lingüístico". *DELTA (Documentação de Estudos em Lingüística Teórica e Aplicada)*. 5:169-184

NARO, Anthony J. & VOTRE, Sebastião J. (1998b) "Discourse motivations for linguistic regularities: verb/subject order in spoken Brazilian Portuguese". *Probus: International Journal of Latin and Romance Linguistics* (to appear).

NARO, Anthony J. & VOTRE, Sebastião J. (1992) "Mecanismos funcionais do uso da língua: função e forma". *DELTA (Documentação de Estudos em Lingüística Teórica e Aplicada).* 8:285-290.

NARO, Anthony J. & PEREIRA SCHERRE, Maria Marta (1993) "Sobre as origens do português popular do Brasil". *DELTA (Documentação de Estudos em Lingüística Teórica e Aplicada).* 9:437-45.

NARO, Anthony J. & PEREIRA SCHERRE, Maria Marta (1997) "Variable Concord in Portuguese: the situation in Brazil and Portugal". In *Current issues in pidgin and creole linguistics* , ed. by John McWhorter. Amsterdam: Benjamins (to appear).

SANKOFF, G. & S. LABERGE (1980) "On the acquisition of native speakers by a language". *The Social Life of Language* 195-210, ed. by. Gillian Sankoff. Philadelphia: University of Pennsylvania Press.