



## Acta Linguistica Hafniensia: International Journal of Linguistics

Publication details, including instructions for authors  
and subscription information:

<http://www.tandfonline.com/loi/salh20>

### Simulating the genesis of Mauritian

Mikael Parkvall<sup>a</sup>, Fredrik Jansson<sup>b</sup> & Pontus Strimling<sup>b</sup>

<sup>a</sup> Department of Linguistics, Stockholm University,  
Stockholm, Sweden

<sup>b</sup> Centre for the study of Cultural Evolution, Stockholm  
University, Stockholm, Sweden

Published online: 13 Aug 2014.

To cite this article: Mikael Parkvall, Fredrik Jansson & Pontus Strimling (2013)  
Simulating the genesis of Mauritian, Acta Linguistica Hafniensia: International Journal of  
Linguistics, 45:2, 265-273, DOI: [10.1080/03740463.2013.900998](https://doi.org/10.1080/03740463.2013.900998)

To link to this article: <http://dx.doi.org/10.1080/03740463.2013.900998>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

## Simulating the genesis of Mauritian

Mikael Parkvall<sup>a\*</sup>, Fredrik Jansson<sup>b</sup> and Pontus Strimling<sup>b</sup>

<sup>a</sup>*Department of Linguistics, Stockholm University, Stockholm, Sweden;* <sup>b</sup>*Centre for the study of Cultural Evolution, Stockholm University, Stockholm, Sweden*

This paper presents a computer simulation of the genesis of Mauritian Creole. The input consists of detailed demographic data and typological information on Mauritian as well as the languages which contributed to its birth. The simulation is deliberately a simplistic one – the idea is to have as few potentially controversial assumptions as possible built into the model, and add additional parameters only to the extent that its output differs from the real-world result. As it turns out, the model generates a language which is highly similar to Mauritian as it is spoken today, and thus, very little “tweaking” seems necessary. Most notably, the model produces the desired result without the postulation of targeted language acquisition, and while one cannot conclude that this was not a part of the creolisation process, our simulation suggests that it is not a necessary assumption.

**Keywords:** creole; creolisation; pidginisation; computer simulation; Mauritius; Mauritian French Creole

### 1. Introduction

Simulations regarding creole genesis are rather scarce, and the only exceptions we are aware of are Nakamura, Hashimoto, and Tojo (2007), Nowak, Komarova, and Niyogi (2001) and Satterfield (2001, 2008). None of them, however, use actual demographic data (at least not from the time and the place where the creoles in question most likely arose).<sup>1</sup>

In this context, the island of Mauritius and the French-lexifier creole spoken there provide a uniquely suitable testing ground, since the language arose on an island with no previous human habitation, and whose early settlement history is documented in detail which is paralleled by no other creole-speaking area (Baker and Corne 1982).

The mathematical details of our simulation are outlined in Jansson, Strimling, and Parkvall (forthcoming), but for the general linguist, it should suffice to know that our virtual settlers (free and enslaved) arrive at the island with no previous

---

\*Corresponding author. Email: [parkvall@ling.su.se](mailto:parkvall@ling.su.se)

<sup>1</sup> An anonymous reviewer claims this is untrue as far as Satterfield’s publications are concerned. However, Satterfield assumes that Sranan, her test case, arose in Surinam, whereas we think there are very good reasons to assume that an early form of it was brought in to that country from elsewhere (Baker 1999; McWhorter 1995; Parkvall 1999). If that is indeed the case, then the relevance of Surinamese obviously decreases drastically.

knowledge of the languages of the other groups. They attempt to communicate by means of linguistic traits found in their respective mother tongues (initially) or which they have encountered on the island (in time). Whenever communication is successful, the material used, even if not part of the speaker's native tongue, is more likely to be chosen in successive communication attempts.

Since we have the "correct answer" in that we know what the result (i.e., Mauritian Creole) is like, we can then compare to what extent the outcome of this simplistic simulation is similar to the one actually documented. The similarities or differences, as the case may be, between the two could potentially teach us something about which additional assumptions need to be made in order to account for the process of creolisation on Mauritius.

The specific questions addressed are the following:

- Did the speakers of the substrate languages strive to create a lingua franca by whatever means available (e.g., Baker 1990, 1995; Smith 2006), or did they specifically attempt to acquire the lexifier (which is an explicit or implicit assumption of the vast majority of creolistic writings)? In the latter scenario, lack of access to the lexifier hindered full acquisition, and instead generated a creole.
- How fast did the creole develop? Suggestions in the literature vary greatly in this regard, from an abrupt genesis to one stretching over several generations.
- Do creoles develop from pidgin languages? From having been uncontroversial, this view of creole genesis has been abandoned by many in recent years.
- To what extent do creole structures derive from the languages in contact? Does everything (or most traits) originate in these, or can linguistic features be present or absent regardless of what is offered by the input languages?

The simulation is expected to create a language with French vocabulary, and with a structure (whether French-like or not) which is common to all its speakers. These two criteria, obviously, are fulfilled by the real-world outcome, and therefore, any simulation results not having these two properties must be ruled out as wrong. After having excluded these, the question is what amount of tweaking – if any – is necessary to obtain a result similar to Mauritian Creole. One possible such adjustment would be to add a factor mimicking targeted<sup>2</sup> second language acquisition (SLA), by which the language creators would be more prone to include elements heard from French-speakers than from others. Should the addition of such a factor not provide a better result, our conclusion

<sup>2</sup> By "targeted SLA" we refer to a process in which learners are actively seeking to acquire an existing language. This differs from pidginisation, where, we believe, the target is simply successful communication, regardless of the etymological composition of the lingua franca.

would be that targeted SLA is not a necessary (albeit still *possible*) assumption in accounting for creole genesis.

## 2. The demographic data

The demographic data used here is derived mainly from Baker and Corne (1982), which covers the first 15 years (1721–1735) of settlement of Mauritius. As is well known to creolists, plantation colonies typically start out with a numerical dominance of Europeans and a small number of others. By time, however, the non-European (typically enslaved) population grows to outnumber the lexifier speakers.

This is true for Mauritius as well, where non-Europeans became more numerous than Frenchmen after a mere nine years, after which the white population stabilised at around a tenth of the total. This demographic disproportion does indeed suggest a limited exposure to the lexifier (French) on the part of the slaves.

The non-Francophone population was brought to the island from a wide variety of locations, for which we have attributed the following likely languages:

The Mozambican Bantu languages, however, only appear on Mauritius after 1735, i.e., right after the end of the period studied by Baker and Corne, and thereby also the end of our simulation.

No assumptions are made with regard to the individual characteristics (age, sex, social standing, etc.) of the virtual creolisers. While this might be perceived as a weakness, we in fact think the simplicity of the model could be interpreted in the opposite way – we want to make the smallest possible number of aprioristic assumptions to begin with, and start adjusting the parameters only if the simplest model fails to generate the right outcome.

Table 1. Inferred languages used on early Mauritius.

Origin	Assumed language
Europe	French
Réunion	French or Bourbonnais <sup>a</sup>
Madagascar	Malagasy
Senegal	Wolof and Manding languages
French India	Tamil and Bengali
Benin	Gbe languages
Mozambique	Makhuwa and related Bantu languages

<sup>a</sup> The neighbouring island of Réunion was known at the time as *Île Bourbon*, and *Bourbonnais* refers to an early variety of Réunionnais Creole. This is also a French-lexicon variety, but very different from Mauritian, and not mutually intelligible with it. Its most characteristic feature is that it is considerably closer to French than Mauritian is. Bourbonnais is what Robert Chaudenson in a large number of publications has argued to be the ancestor of Mauritian. Apart from the fact that this, in our view, has been resoundingly disproved (primarily through Baker and Corne 1982), it is unclear to what extent Bourbonnais actually differed from French at the time. Because of this (and there not being a description of Bourbonnais in the first place), Réunionnais immigrants are treated as speakers of French in our simulations.

### 3. The linguistic data

The linguistic component of our simulation consists of features from the *World Atlas of Language Structures* (WALS) (Haspelmath et al. 2005), combined with the structural features for creoles of Grant and Baker (2007, 24–27) and the *UCLA Phonological Segment Inventory Database* (UPSID). These sources all deal with linguistic features in terms of whether a given trait is present or not in a given language.

Since the languages just enumerated (and Mauritian Creole) are not always present in these sources, a variety of reference grammars were used to fill in the blanks. We finally decided to use only those features which were available for all languages involved, which resulted in a database consisting of 128 features (89 phonological, 30 syntactical and 9 morphological; the complete list can be found in Jansson, Strimling, and Parkvall (forthcoming)).

In accordance with our search for maximal simplicity, there is no weighting of any kind of the features.

### 4. Running the simulation

As mentioned above, the details of our algorithms can be found in Jansson, Strimling, and Parkvall (forthcoming).

At the outset, each individual knows only their native tongue. Many people are of course likely to have known more than one language, but the geographical diversity in this case was such that most second language competence that people had on arrival would have been of little use (i.e., knowing several West African instead of just one wouldn't have been particularly beneficial when communicating with an Indian or a Malagasy).

The individuals arrive (and leave!) the community according to the actual historical data.<sup>3</sup> Once present, the individuals try to communicate in pairwise interactions. If understanding is achieved, the likelihood of the interlocutors trying the same strategy in the next conversation increases.

A simulation result is only considered successful if it produces a (1) common language with (2) predominantly French vocabulary.

There is an interesting difference between lexicon-based and structural features which turns up in our results. In line with the Saussurean “arbitrariness of the linguistic sign”, vocabulary items can take virtually any shape, and a given word is unlikely to have the same realisation in two unrelated languages. Syntactical properties, however, have fewer parameter settings, and this leads to accidental matches. To take the simplest possible example: if we have more than six languages present (which is true in this case), at least two of them are bound to share the same basic word order, since there are only six options to choose from (unless we count free word order as a seventh possibility). Therefore, languages

<sup>3</sup> We experimented also with adding births and deaths, but this proved to have a negligible effect on the outcome.

Table 2. Similarities between (modern) Mauritian, its input languages, and our simulation.

	All 119 features	Phonology	89 features	Syntax	30 features
SIMULATION	92%	SIMULATION	93%	SIMULATION	87%
French	84%	Malagasy	87%	French	87%
Malagasy	82%	French	83%	Wolof	80%
Wolof	77%	Bengali	81%	Malagasy	70%
Bengali	72%	Manding	81%	Gbe	67%
Gbe	71%	Wolof	76%	Bengali	47%
Manding	71%	Tamil	73%	Manding	43%
Tamil	63%	Gbe	73%	Tamil	33%

are capable of “teaming up” and assist each other in the propagation of syntactical features, which is virtually impossible when it comes to lexical material.

Excluding morphology for the time being, the best of our simulation runs produced a language which behaved like (modern) Mauritian for 92% of the features investigated.

The similarities vis-à-vis the other languages involved are as set out in Table 2.

It may be noted that our simulation comes closer to actual Mauritian than does any one of the input languages, which is interesting in itself. French is the language closest to Mauritian with not only a close to 100% resemblance in the vocabulary, but also with an 87% resemblance for syntax, while the phonology is closest to Malagasy.

We also made a comparison with languages not involved (recall that one of our original questions was “To what extent do creole structures derive from the languages in contact?”).

Limiting ourselves here to the features found in WALS, the input languages match Mauritian in between 93% (French) and 39% (Tamil) of the cases, with an average of 63%. The languages of the world in general typically fare less well. Not by a great margin, but still a statistically significant one.

## 5. Failed predictions

As already noted, the results of our simulation are highly similar to actual Mauritian, but for fifteen features, there is a mismatch between the simulation results and the actual state of affairs. The predicted features are set out in Table 3.<sup>4</sup>

In our view, nine of these 15 can be explained by the pidgin past of Mauritian Creole. The most well known and least controversial feature of pidgins is their complete or near-complete lack of morphology and dearth of grammatical elements in general. Because of this, we would not expect a pidgin ancestor of Mauritian to have had a morphological imperative, extensive suffixing or grammatical gender regardless of the preferences among the input languages.

<sup>4</sup> For the definition of the WALS features, please refer to that publication (Haspelmath et al. 2005).

Table 3. Features with regard to which the simulation made incorrect predictions.

absence of /ɣ/ and /h/
presence of /ɔ/, /ɛ/, /ɲ/ and /ʃ/
Number of Genders, Suffixed nominal plural, Suffixing inflexional morphology, Morphological Imperative
Nominal and Locational Predication (two features), Nonverbal adjectives, Preverbal negation
Postposed definite article

A few additional features do not necessarily involve morphology, but are still in line with characteristics typical of pidgins. The feature “verbal encoding of predicative adjectives”, for instance, follows from the absence of an equative copula (in combination with the lack of morphology). This automatically yields a situation whereby adjectives (*You Ø sick*) are indistinguishable (or virtually so) from verbs (*You run*). We have also included the use of a simple preverbal negator in this category – while not a by-product of analyticity, it does represent a strategy typical of pidgins.

Some morphology is indeed present in modern Mauritian, but this appears to represent later developments, that is, it emerged long after the period that concerns us here (see e.g., Baker 2009 and Guillemin 2009). At that time, Mauritian had had native speakers for a considerable time, and these later developments are thus most likely unrelated to the languages present in the initial contact situation.

Apart from the pidgin-related traits, there remains the question of the phonemes /h/, /ɣ/ and /ʃ/, /ɲ/, /ɛ/, /ɔ/. The former two were predicted to be absent, but in fact occur in Mauritian, whereas the opposite is true for the other four.

These four are such that they all occur in French, but failed to make their way into the creole. They are all absent from Malagasy, speakers of which dominated the non-European population at times (at the beginning of the fifteen-year period studied here, but also after it). In fact, a French phonology “filtered” through a Malagasy one would come closer to Mauritian phonology than either our simulation or any of the languages involved.

The remaining two phonemes, /h/ and /ɣ/, are more marginal failures. The latter corresponds to French /ʁ/, and is articulatorily close to it, and especially to some of its dialectal French realisations. /h/ is lacking from European French, but once was present, and still is in some of its dialects (including several varieties spoken overseas). The fact that the phoneme occurs in Mauritian (other than in loanwords) in precisely those positions where standard French had it proves, of course, that it was in fact part of the lexifier input of Mauritian.

Thus, the features which our model failed to predict correctly can be explained if we make two additional assumptions: (1) Mauritian started its life as a pidgin (entirely or virtually) devoid of morphology, and (2) the phoneme inventory crystallised at an early date, when the non-European population of the island was dominated by speakers of Malagasy. However, after a period of



Table 4. Attempting to explain the incorrect predictions.

/ɣ/, /h/	Not really wrong
/ɔ/, /ɛ/, /ɹ/, /ʃ/	Adaptation to Malagasy phonology
Number of Genders, Suffixed	Morphology, and loss is
nominal plural, Suffixing inflexional	therefore expected due to
morphology, Morphological Imperative	pidginisation
Nominal and Locational Predication,	Other expected consequences of
Nonverbal adjectives, Preverbal negation	pidginisation
Postposed definiteness article	Later development (c. 1820, according to Guillemín 2009).

demographic decline, Malagasies again dominated the slave population around the middle of the eighteenth century (Grant and Baker 2007, 202), which opens up the possibility of the phonology gelling at this point in time.

The failed predictions mentioned above can therefore be summarised and categorised as in Table 4.

Thus, most of the failed predictions are such only if we assume that Mauritian has not changed since its birth (obviously an absurd claim), and that it does not have a pidgin past.

6. Conclusions

The model used to simulate the birth of Mauritian was deliberately simplistic. By not including additional (and potentially controversial) factors, we can then see which kinds of fine-tuning are needed to produce the observed result (i.e., Mauritian as later documented).

One of the questions asked at the beginning of this paper was whether creolisation is the result of (targeted) SLA, or simply the outcome of communication attempts without the people involved striving to acquire a specific pre-existing language. The high similarity of our simulation results to the actual language would seem to indicate that the identification of a “target language” is not a necessary assumption. However, could it be that we would have fared even better by including such a factor? The answer is no. In the simulation, the most straightforward way to include this would be to make the agents more prone to accept a certain input if it is produced by a French-speaker. Now, recall that almost all of the vocabulary already is of French origin, just as it is in actual Mauritian. The only effect of a learning bias favouring French would be to make the grammar (and the phonology) more like French, which is clearly not in accordance with the empirical facts. None of our failed predictions would become correct if brought more in line with French, and so, the result of the simulation would not only not improve, but rather distance itself from Mauritian if targeted SLA were introduced.

As for the other three questions, we have shown that assuming a pidgin stage in the history of Mauritian explains the vast majority of the seemingly incorrect

predictions made by the model. In our view, this makes the traditional (but nowadays questioned) pidgin-to-creole scenario a likely one. It might also be added that there is no other process of language change known to completely eradicate an entire morphological machinery from a language in such a short time span.<sup>5</sup>

Something can also be said about the pace at which a creole develops – the fact that we only took the first 15 years into account<sup>6</sup> and still got a near-perfect (if pidginisation is taken into account) result makes it tempting to assume that a longer period of time is not required.

Finally, to what extent do creole structures derive from the languages in contact? Here, there is a marked difference between morphology and the other components. In our simulation, only the languages known by the immigrants were taken into account, and there was no mechanism beyond demographics to filter these, much less a component to simulate spontaneous drift. Yet, we obtained a near complete match between the simulation result and contemporary Mauritian with regard to the non-morphological features provided. For morphology, we would of course not expect much matching if we accept a pidgin past. But for syntax and phonology, it does seem that the outcome does derive rather straightforwardly from the multilingual input – recall that the languages of the world as a whole match Mauritian (actual and simulated) less well than do the ones which were present on the island.

Thus, our findings indicate that:

- It is not necessary to assume an acquisitional bias towards the lexifier language (i.e., targeted learning)
- Creoles develop quickly
- They do so from pidgin languages
- The input languages do to a great extent determine the phonological and syntactic make-up of the language, but have little or no influence on its morphology.

Those four conclusions seem to hold in the case of Mauritian. While we believe they would also be valid for other creologenetic settings, that remains to be shown. We would of course like to test this, but a problem is that few other locations are blessed with demographic data of anywhere near the quality that Baker and Corne (1982) provided for Mauritius.

### Notes on contributors

Mikael Parkvall is active at the department of linguistics of Stockholm University, and has long experience of research on Creoles and other contact languages. Pontus Strimling and Fredrik Jansson are both mathematicians based in Stockholm, the former at the Centre for the study of Cultural Evolution, and the latter at the Institute for Futures Studies.

<sup>5</sup> Modern Mauritian does have a few affixes (e.g., diminutive *ti-*, and a verbal suffix *-e* used in certain syntactical positions), but none is directly inherited from its lexifier.

<sup>6</sup> The 15-year limit was imposed by the source, and not chosen by us.

Fredrik Jansson is a postdoctoral researcher at the Institute for Futures Studies in Stockholm and affiliated with the Centre for the Study of Cultural Evolution at Stockholm University.

Pontus Strimling is a research fellow at the Institute for Futures Studies in Stockholm and a guest researcher at the Centre for the Study of Cultural Evolution at Stockholm University.

## References

- Baker, Philip. 1990. "Off Target?" *Journal of Pidgin and Creole Languages* 5 (1): 107–119.
- Baker, Philip. 1995. "Motivation in Creole Genesis." In *From Contact to Creole and Beyond*, edited by Philip Baker, 3–15. London: University of Westminster Press.
- Baker, Philip. 1999. "Investigating the Origin and Diffusion of Shared Features Among the Atlantic English Creoles." In *St. Kitts and the Atlantic Creoles*, edited by Philip Baker, and Adrienne Bruyn, 315–364. London: Westminster University Press.
- Baker, Philip. 2009. "Productive Bimorphemic Structures and the Concept of Gradual Creolisation." In *Gradual Creolization*, edited by Rachel Selbach, Hugo Cardoso, and Margot van den Berg, 27–53. Amsterdam: John Benjamins.
- Baker, Philip, and Chris Corne. 1982. *Isle de France Creoles: Affinities and Origins*. Ann Arbor: Karoma Publishers.
- Grant, Anthony, and Philip Baker. 2007. "Comparative Creole Typology and the Search for the Sources of Mauritian Creole Features." In *The making of Mauritian Creole*, edited by Philip Baker, and Guillaume Fon Sing, 197–216. London: Battlebridge Publications.
- Guillemin, Diana. 2009. "The Mauritian Creole Determiner System: A Historical Overview." In *Complex Processes in New Languages*, edited by Enoch Aboh, and Norval Smith, Amsterdam & Philadelphia: John Benjamins. 173–200.
- Haspelmath, Martin, Matthew Dryer, David Gil, and Bernard Comrie, eds. 2005. *The World Atlas of Language Structures*. Oxford: Oxford University Press.
- Jansson, Parkvall, and Strimling. In press. "Modelling the Evolution of Creoles" *Language Dynamics and Change*.
- McWhorter, John. 1995. "Sisters Under the Skin: A Case for Genetic Relationship between the Atlantic English-based Creoles." *Journal of Pidgin and Creole Languages* 10 (2): 289–333.
- Nakamura, Makoto, Takashi Hashimoto, and Satoshi Tojo. 2007. "Exposure Dependent Creolization in Language Dynamics Equation." In *JSAI 2003/2004, LNAI 3609*, edited by Akito Sakurai, Kôiti Hasida, and Katsumi Nitta, 295–304. Berlin: Springer-Verlag.
- Nowak, Martin A., Natalia L. Komarova, and Partha Niyogi. 2001. "Evolution of Universal Grammar." *Science* 291 (5501): 114–118.
- Parkvall, Mikael. 1999. "Feature Selection and Genetic Relationships Among Atlantic Creoles." In *Spreading the Word*, edited by Magnus Huber, and Mikael Parkvall, 29–66. London: Westminster University Press.
- Satterfield, Teresa. 2001. "Toward a Socio-Genetic Solution: Examining Language Formation Processes through SWARM Modeling." *Social Science Computing Review* 19 (3): 281–295.
- Satterfield, Teresa. 2008. "Back to Nature or Nurture: Using Computer Models in Creole Genesis Research." In *Variation, Selection Development: Probing the Evolutionary Model of Language Change*, edited by Regine Eckardt, G. Jäger, and Tonjes Veenstra, 143–178. Amsterdam: Mouton de Gruyter.
- Smith, Norval. 2006. "Very Rapid Creolization in the Framework of the Restricted Motivation Hypothesis." *Language Acquisition and Language Disorders* 42: 49–65.