The vitality or endangerment of some nonindigenous languages:
A response to Mufwene

NALA H. LEE
National University of Singapore

One critical assumption that Salikoko Mufwene (2017) makes about the field of language endangerment and loss is that linguists engaged in language endangerment, documentation, and revitalization are concerned with indigenous languages, which naturally leaves out nonindigenous languages. This response concerns itself with addressing this assumption, with a focus on a particular group of nonindigenous languages. It provides insight on the levels of endangerment of pidgins, creoles, and mixed languages for which we have information, and considers some reasons why it is important to focus on the endangerment and loss of these types of nonindigenous languages.

Keywords: language endangerment, vitality, language endangerment index, contact languages, pidgins, creoles, mixed languages

1. Introduction. For various reasons, some of which are addressed here, it is often assumed that the field of language endangerment and loss (LEL) is concerned predominantly with indigenous languages. Salikoko Mufwene (2017) raises a series of interesting issues for the field, chief among which is the observation that the discourse leaves out ‘nonindigenous’ populations, such as in the ‘European settlement colonies’, ‘Bantu settled territories’, and ‘plantation settlement colonies’ (pp. e212–e213). This raises the noteworthy question of whether there is an overreliance on the terminology ‘indigenous’ among those engaged in LEL, and whether the conversation can and should include nonindigenous languages.

In particular, I here focus on pidgins, creoles, and mixed languages, even though most of these have comparably shallow time depths when positioned against early European settlements. While Mufwene states that the ‘typical approach to LEL in linguistics has suffered from lack of historical time depth, despite passing allusions to language loss in Europe before the fifteenth century’ (p. e203), there are possibly good reasons for why this is so. The field of LEL took off mainly in the very late twentieth century, as a response to observations such as Krauss’s (1992) that at least half and possibly as many as 90% of the world’s languages will become extinct or doomed by the end of the present century. The expansion of the field has also been aided in no small part by technological advancements. Linguists are far better enabled than ever to document and revitalize languages at rates unlike before, with software tools that aid transcription and translation, to name just two of the many processes involved. They are also far better positioned to map and trace language endangerment patterns around the world on platforms such as the Endangered Languages Project portal, which features the Catalogue of Endangered Languages (ELCat; http://www.endangeredlanguages.com). While we know that languages have indeed come and gone in the past, we now possess tools to tell us at what rate. For example, using the LANGUAGE ENDANGERMENT INDEX (LEI; Lee & Van Way 2016), ELCat shows that one language becomes extinct every three months (see Campbell et al. 2013). This knowledge comes with its own set of obligations. As scientists, those in this field understand the need to preserve linguistic diversity, so that a natural laboratory of 7,000 or so languages (Evans & Levinson 2009) is not lost. As scientists who understand that their work can potentially benefit communities of speakers, some are obliged to provide assistance when community
members themselves ask for help. A good number of researchers who engage in document-
mentation and revitalization are aware of the rate of language endangerment and the
various reasons why it is necessary to do something about it right now more so than
ever, before more is lost.

The core approach to LEL is hence necessarily synchronic rather than diachronic. I
believe that a number of the responses to Mufwene 2017 will detail the ways in which
scholars engaged in the field are doing valuable work, so I focus here instead on clarifi-
ing the notion that the term ‘indigenous’ is overemphasized in LEL by looking at a
particular group of nonindigenous languages with much shallower time depths, but
which currently still concern us.

It is clear why the field of LEL may be perceived as being concerned primarily with
the indigenous. Mufwene notes that it is a term that the United Nations and UNESCO
have adopted in their discourse on LEL. He goes on to problematize the term, question-
ing if it does not simply represent a divide between ‘Europe, where there are suppos-
edly no indigenous peoples’ (p. e212) and ‘the rest of the world, especially the former
European colonies, inhabited first by indigenous people’ (p. e212), and if there is not ‘a
part of the world where people of European descent may be considered indigenous’ (p.
e212). If so, Mufwene is justified in stating that ‘the absolute usage of the term has im-
lications for our engagement as linguists with explaining how and why some lan-
guages die, while others are maintained and sometimes spread demographically (and
geographically)’ (p. e212). The issue then lies in whether the term is used in an absolute
sense in LEL. UNESCO does indeed appear to emphasize ‘indigenous’, with one if its
cultural themes being ‘[m]aintaining indigenous languages, conserving biodiversity’
(UNESCO 2017), and there is no question that there are links between indigenous tra-
ditions of classification and knowledge of biodiversity. While this sort of knowledge is
unquestionably important, those involved in LEL are also aware of the many other rea-
sons why language loss is alarming, such as the fact that language loss compromises the
ability of linguists to understand the full range of what is possible in human language
and cognition (Lee & Van Way 2016). In that same vein, researchers and agencies in-
volved in LEL represent a broad range of concerns and are interested in the study of all
endangered languages, so whether these languages are indigenous in the sense utilized
by UNESCO does not quite matter.

Within the field of LEL, the inclusion of nonindigenous languages is evident in several
ways. First, except for UNESCO, major agencies involved in LEL do not differentiate be-
tween indigenous languages and nonindigenous ones. For example, the Endangered Lan-
guages Project (ELP) online platform (http://www.endangeredlanguages.com), which
provides information on over 3,000 endangered languages around the world, states that it
is ‘[a] worldwide collaboration to strengthen endangered languages’. Similarly, the En-
derangered Languages Fund (http://www.endangeredlanguagesfund.org) states that it is
aimed at ‘[s]upporting the documentation and revitalization of the world’s endangered
languages’. Likewise, the Endangered Languages Documentation Programme (ELDP;
http://www.eldp.net), housed at SOAS, states that its mission is ‘the preservation of en-
derangered languages globally’. All of these agencies value work on all types of endangered
languages, and whether the languages are indigenous is secondary—the word ‘indege-
nous’ does not occur on any of the agencies’ main webpages.

Aside from the statements of purpose, it is also useful to look at the types of lan-
guages and projects that these agencies feature. In addition to what would constitute
conventionally indigenous languages, such as Akuntsú (Aragon 2014), a Tupian lan-
guage spoken by survivors of a genocide in the southeastern region of Rondônia,
Brazil, languages that are covered by these agencies include those that are not indigenous in the same sense, such as Michif, an exemplar mixed language based on French and Cree spoken in the United States and Canada (Bakker 1997; see Sammons 2011). Other traditionally nonindigenous languages covered by projects under the ELDP include creoles such as Malaccan Creole Portuguese (Pillai 2011), which has a Malay substrate and Portuguese lexifier, and sign languages such as Cameroon Sign Language (Lutalo-Kiingi 2012), which is a language that has been influenced by American Sign Language and Langue de Signes Française. None of these languages are strictly related to an indigenous culture, but they are clearly valuable to researchers engaged in language documentation, as well, of course, to the populations that use these languages. Hence, while it is perhaps logical that UNESCO, being the major organization that it is, may be taken to represent the dominant view on LEL, it is important to consider the perspectives of these other agencies and researchers who have been working firsthand with speakers of nonindigenous and indigenous languages.

As just demonstrated, most agencies and researchers who have a stake in the field of LEL would not deem the study of indigenous languages to be more vital than that of nonindigenous languages. That being said, Mufwene’s article does raise the concern that not enough is being said about anything that falls outside the domain of conventional indigenous languages, especially if the public perception of the field is informed by UNESCO’s focus, which puts the spotlight on indigenous languages. Crucially, this concern that more should be said about nonindigenous languages is not a new one. Garrett (2006:177) notes in particular that contact languages are hardly mentioned or focused upon in discussions of language endangerment, asking ‘why should the precarious state of so many languages go virtually ignored in discussions of language endangerment?’.

This paper takes the opportunity to highlight the state of endangerment or vitality of this group of languages, conventionally deemed to be nonindigenous, including pidgins, creoles, and mixed languages. Other nonindigenous types of languages, such as sign languages, should also not be left out of the LEL discussion, but they are not the focus of this response. A general tool of assessing language endangerment is applied to these languages, drawing attention to how these assessment methods can be applied to nonindigenous languages (in addition to indigenous languages). This is followed by a discussion underscoring why linguists should be concerned about the endangerment and loss of these nonindigenous languages.

2. On the state of pidgins, creoles, and mixed languages. The contact languages surveyed here are assessed on the LEI, developed by Lee and Van Way (2016) for use with ELCat. EL Cat, hosted on the ELP portal as mentioned above, does feature endangered contact languages and their levels of endangerment, but the list is not exhaustive. Twenty-four pidgins and creoles are listed on the site: Barikanchi, Berbice Creole Dutch, Chinook Wawa, Gibanawa, Iha Based Pidgin, Javindo, Kodiak Russian Creole, Louisiana Creole, Malabar Sri-Lanka Portuguese, Malaccan Creole Portuguese, Nauru Pacific Pidgin, Ndyuka-Trio Pidgin, Ngatik Men’s Creole, Nicaragua Creole English, Nubi, Palenquero, Macao Creole Portuguese, Petjo, Pitcairn-Norfolk, San Miguel Creole French, Settla, Sri Lanka Malay, Torres Strait Creole, and Yilan Creole. The six mixed languages featured on the site are Copper Island Aleut, Kallawayá, Mbugu, Michif, Tagdal, and Wutunhua. There are a good number of other contact languages whose vitality or endangerment should be considered. For the purpose of more fully understanding the state of contact languages around the world, I also assess languages that are featured in the Atlas of pidgin and creole language structures online (APiCS; http://apics-online.info/) for their individual levels of endangerment, bringing the total
number of languages surveyed to ninety-eight. The information that is available in APiCS for each language, including the language’s sociolinguistic features, allows for an assessment of endangerment using the LEI.

The mechanisms of the LEI (Lee & Van Way 2016) are beyond the bounds of this paper, but for our present purposes, the LEI can be understood as a system of assessing language endangerment built on four factors that are easily comparable and more readily available than others: intergenerational transmission (whether the language is being transmitted to the children), absolute number of speakers, speaker number trends (whether stable, increasing, or decreasing), and domains of use (whether the language is used in official domains or unofficial ones). Note that population size (brought up in the discussion by Mufwene and said to be an inadequate explanation of LEL) is but one of the factors to be considered. It is not the most important factor, but it is necessary, especially since there are languages about which nothing is known except for the number of speakers. Intergenerational transmission is doubly weighted in the assessment, for there is clearly no continuity of the language without it. All languages can then be assessed objectively for a level of endangerment, regardless of how much or how little information there is on a given language (unless no information is available at all on any of the factors). The more factors are used in the assessment, the more certain one can be about how endangered a language is, and an assessment that utilizes fewer factors is one that should be regarded as having less certainty. For a more technical discussion of how the LEI is applied, see Lee & Van Way 2016. The levels of endangerment, from least to most endangered, are: vulnerable, threatened, endangered, severely endangered, and critically endangered. At the two ends of the scale are also safe languages and dormant languages (languages that are thought to have lost their last remaining speakers within approximately the last fifty years). It is important to note that LEI errs on the side of caution, putting more languages in endangered than safe categories, given the magnitude of the problem of language endangerment and the alarmingly fast rate at which the world’s linguistic diversity is being lost.

Using the exact method utilized by ELCat, the levels of endangerment (or vitality) of the ninety-eight contact languages around the world were established. A total of twenty pidgins, seventy creoles, and eight mixed languages are represented, and specific results are reported in Lee 2017. General results are reported here to give a sense of global trends: of the ninety-eight languages, ten are dormant: Ndyuka-Trio Pidgin, Chinese Pidgin English, Chinese Pidgin Russian, Pidgin Hawaiian, Eskimo Pidgin, Bungi, Negerhollands, Batavia Creole, Berbice Creole Dutch, and Copper Island Aleut. Another eleven languages are critically endangered, these being Chinook Wawa, Yimas-Arafundi Pidgin, Baba Malay, Javindo, Kodiak Russian Creole, Macao Portuguese Creole, Petjo, San Miguel Creole French, Principense, Malabar-Sri Lanka Portuguese, and Michif. Five languages are severely endangered: Singapore Bazaar Malay, Louisiana Creole, Ngatik Men’s Creole, Cavite Chabacano, and Kallawaya. Seven are endangered: Malaccan Creole Portuguese, Pitcairn-Norfolk, Sri Lanka Malay, Yilan Creole, Diu Indo-Portuguese, Ternate Chabacano, and Wutunhua. The only contact language that can be deemed ‘safe’—due to having at least 100,000 speakers, stable or increasing numbers, and usage in all domains—is Seychelles Creole (for which information on three factors was used, since information about the doubly weighted factor of intergenerational transmission was not available from the sources; the level of certainty of this assessment stands at 60%). Other languages that are also designated as national, such as Haitian Creole, Tok Pisin, and Bislama, all have some official domain in which their use is limited, for example, in education or in the courts. These languages thus fall within the vulnerable category. A total of sixty-one pidgins, creoles, and mixed languages fall into the
threatened and vulnerable categories. Note also that there are contact languages for which there is no vitality information. These include the pidgins Barikanchi, Gibanawa, and Settla.

What do the numbers portend for this particular group of nonindigenous languages? Essentially, the number of pidgins, creoles, and mixed languages that can be deemed safe form an extremely small fraction of the total number of languages surveyed at 1.0% (one out of ninety-eight languages). A total of 10.2% (ten of ninety-eight) are dormant. Another 11.2% (eleven of ninety-eight) are critically endangered, while 5.1% (five of ninety-eight) are severely endangered, and 7.1% (seven of ninety-eight) are endangered. Another 22.4% (twenty-two of ninety-eight) are threatened, and 39.8% (thirty-nine of ninety-eight) are assessed to be vulnerable. Nothing is known about the vitality of the remaining 3.1% (three of ninety-eight) of the languages surveyed.

In all, the proportion of pidgins, creoles, and mixed languages that are at some level of risk is high at 85.7%. This proportion is almost double that of the languages around the world that are at a similar level of risk (including both traditionally indigenous and nonindigenous languages like the contact ones discussed here). The ELP attempts to feature all of the known endangered languages (with different levels of endangerment or vitality), and it includes a total of 3,153 languages. If there are indeed 7,099 languages in the world (Simons & Fennig 2017), this would mean that 44.41% of the world’s languages are at some level of risk. As the numbers have it, the risk of endangerment for pidgins, creoles, and mixed languages is thus almost twice as great as that for all of the world’s languages. The rate of endangerment and loss of languages around the world is worrying for its threat to linguistic diversity, which would also compromise the abilities of linguists to study what human cognition is capable of, but it is also worrying for reasons having to do with the identity and well-being of speakers. If the numbers are anything to go by, researchers should be paying significant attention to the risk of loss faced by speakers of pidgins, creoles, and mixed languages.

3. On why pidgins, creoles, and mixed languages matter in the endangerment discussion. It is perhaps not surprising that a large number of these contact languages are particularly vulnerable. Frank (2007:1) discusses the stigma associated with creoles in particular, highlighting that it is really ‘an old and often-repeated story’ that ‘creole languages are popularly derided as corrupt and inferior forms of a “standard” language like English or French’. He uses the examples of St. Lucian Creole and Gullah, noting that ‘[f]or a long time, St. Lucian Creole has been called “broken French” and Gullah and other English Creoles have been called “broken English”’. These attitudes toward creoles and other contact languages inadvertently bear on factors associated with language endangerment, such as speaker number trends and domains of use. Since many of these contact languages emerged in the context of colonial expansion, displacement, and relocation, they usually developed alongside a more dominant language (usually the lexifier), and a good number of them still exist alongside a more dominant language, which may or may not be its lexifier (Garrett 2006). The contact language then often functions as an unofficial language in a diglossic type of relationship with the official language. In relation to creoles being used alongside their lexifiers, Valdman (1987:107) states: ‘Creoles are inexorably destined to dissolve in these major languages via the process of decreolization’ (see also DeGraff 2001).

Singlish is an example of a creole that still exists alongside its lexifier language, English. Spoken in Singapore, the creole is officially discouraged by the government via an ongoing Speak Good English Movement (Wee 2010). Singlish is used mostly in un-
official domains, with official functions such as administration and education being borne by English. An example of a creole that became dormant, as its official functions within the community gradually came to be subsumed by English (which is not its lexifier), is Negerhollands. Once spoken on the islands of St. Thomas, St. John, and St. Croix (the current US Virgin Islands), the Dutch-based creole lost its last speaker in 1987 (van Sluijs 2013).

Given the historical and social nature of contact languages, they are less often used in all domains, though there are clear exceptions such as Seychelles Creole. While it seems strange to consider Singlish as being at risk, the fact of the matter is that a language that is not used in all domains can never be truly safe, and if anything, it is much easier for a language to lose domains than it is to gain them. In South Africa, as of April 2017, English was declared the official language of the courts, with Afrikaans being deemed no longer acceptable for official purposes (Nombembe 2017). While the classification of Afrikaans as a creole has been debated to some extent (see Markey 1982, Thomason & Kaufman 1988), it is generally recognized as a contact language and has been included in APiCS (see den Besten & Biberauer 2013). It is also notable that even creoles that are recognized as national languages seldom have equal and unfettered use in all domains. Haitian Creole is seldom used in the courts (DeGraff 2017), for example, and Tok Pisin is only used in the first three years of formal education if the community has chosen it as a medium of education (Smith & Siegel 2013). Intrinsically, then, contact languages such as creoles, with their use in curbed domains, are by nature highly susceptible to LEL.

Why then should researchers care about the loss of this group of nonindigenous languages, comprising pidgins, creoles, and mixed languages? Is language shift necessarily maladaptive, as Mufwene puts it? And is language shift necessarily maladaptive in the case of pidgins, creoles, and mixed languages, especially since they can be interpreted as adaptive forms of languages that have emerged from the adaptation of particular languages to the requirements of specific language ecologies? In a sense, it is language contact that leads to the emergence of these contact languages, just as it is language contact that leads to language death. Mufwene makes an interesting point, that ‘the study of LEL must certainly be connected to genetic linguistics, which shows language speciation, thus an increase in the number of languages, whereas the literature on LEL has been forecasting a drastic and rapid decrease in the number of languages around the world’ (p. e219). So the question might be: What if a language dies? Another language might be created. Is language death due to language contact part of a normality that is to be left unfought? If anything, no. To demonstrate, if any language covered here in this response were to be lost, there would be significant consequences. The death of any language, even one formed by language contact, denotes loss on several levels. Some of the consequences of language loss include (i) the loss of cultural or ethnic identity (Tsunoda 2005), (ii) the loss of part of the sum of human knowledge (Hale 1992, Crystal 2000), (iii) the loss of linguistic diversity (Hale 1992), and the (iv) loss of languages themselves, on which depends the ability of linguists to discover the full range of what is possible in human language and cognition (Lee & Van Way 2016).

With regard to (i) cultural or ethnic identity, while enslaved or contract laborers gave up their languages in a shift toward a dominant European language, as in the case of plantation settlements, what resulted are creoles that gradually became a part of the speakers’ social identities, such as in the case of Hawai’i Creole English (also called ‘Pidgin’). This language is emblematic of solidarity and local culture (Drager 2012), as are other creoles, mixed languages, and some pidgins. In an interview with BBC News
that highlighted a Kristang (also known as Malaccan Creole Portuguese) revitalization project in Singapore, an eighty-year-old Portuguese Eurasian said, ‘It never dawned upon me that Kristang was my language, I’d always thought it was English … I’ve missed it’ (Wong 2017). Languages, whether indigenous or not, clearly have bearing on one’s cultural and ethnic identity, and even on a sense of well-being that is linked to this sort of identity. That being said, there is seldom a one-size-fits-all solution, and it is important to note that linguists engaged in the field of LEL would never deem it appropriate to proselytize that people should speak the language of the group they were born into, if for whatever reason this is not the wish of the community.

With regard to (ii) language being part of the sum of human knowledge, there is much cultural knowledge that is lost when languages are lost, again regardless of whether the language is indigenous. Kallawaya, spoken in the highlands of Bolivia, is a severely endangered mixed language with a Quechua grammatical base and a varied lexical base that derives partly from Puquina and from other languages in the region (Muysken 2012). Used by healers, the language encodes some 980 species of medicinal plants (UNESCO 2008). Much of this knowledge will be lost if the language is not documented. Aside from botany, other contact languages also come with their own cultural repertoires. In Baba Malay, a critically endangered Malay-based creole with a Southern Min substrate (see Lee 2014), many idioms and sayings are unique to it and not found directly in either of its component languages. For example, *sian pukol tambor lagik ada salah* (literally: ‘deities strike drum also possess mistakes’) means ‘even the deities err, so it is human to make mistakes’. With the loss of languages such as these (as with traditionally indigenous languages), the world loses knowledge and ways in which to look at itself.

The (iii) loss of linguistic diversity and (iv) the loss of languages themselves, which compromises the abilities of linguists to understand the full range of what is possible in human cognition, are interrelated. As Evans and Levinson (2009) put it, the world’s 7,000 or so diverse languages are a natural laboratory. Just as languages such as Yélî Dnye support exciting possibilities for coarticulation, and languages such as Tzeltal show that positionals may be a valid word class (Evans & Levinson 2009), there are many perspectives that can be gleaned from pidgins, creoles, and mixed languages. Each contact language in itself is an experiment. What are all the possibilities of contact-induced change? What happens, for example, when a tone language is mixed with a non-tone language, as in the case of Baba Malay? What happens when a language with clearer word classes is mixed with a language with less clear word classes, as in the case of Kristang? How are complex kinship systems maintained, adapted, or expanded when at least one of the component languages has a complex kinship system, as in the case of Wutunhua, which derives most of its lexical items from Chinese (Sandman 2016)? Is it always a case of simplification? Or are there other outcomes? If nothing is done about these contact languages, it is possible that linguists might simply never know, since those mentioned are all endangered. Pidgins, creoles, and mixed languages provide a whole other depth to the issue of linguistic diversity and should not be ignored, especially when many of them will face the same fate as languages conventionally described as being ‘indigenous’ if nothing is done.

By focusing on the vitality or endangerment of pidgins, creoles, and mixed languages, the discussion here has hopefully brought to the forefront some connections between language loss and language speciation. To conclude, rather than ‘omit the big picture, viz., the coexistence of languages in particular polities and how the competition that sometimes arises is resolved’ (as Mufwene puts it; p. e202), a discussion centered on terms such as
‘language endangerment’, ‘endangered languages’, ‘language loss’, and ‘language death’ highlights the race against time and the potential consequences of loss for linguists and community members. Whether the term ‘indigenous’ is invoked is secondary.

REFERENCES


LEE, NALA H. 2017. The levels of endangerment of pidgens, creoles, and mixed languages around the world. Singapore: National University of Singapore, MS.


[ellhln@nus.edu.sg] [Received 1 June 2017; accepted pending revisions 3 June 2017; accepted 7 June 2017]