1 Introduction

You are sitting in the tropical garden of a café nursing your thirst with a cold beer when you hear the following conversation:

- *Eh41 yu got watch PCK las nait or not?*
- *Ya lor55. Wa lau he kena sen English class! His English so lau ya one meh55?*
- *Ya wat21! Gamen say mas spik gud English ma21. Don play-play! Oderwise people ting yu dam chingchong la21.*[1]

You understand some of the words, but not all of them, because, although they sound English, some are pronounced in 'strange' ways. More importantly, you cannot quite make sense of all that is being said. So was it really English you heard or did you just imagine it? If it was, what was going on with the grammar? Because, surely, there was something wrong somewhere. And what about those little exclamations at the end of each sentence with such strange intonation? You might think that you have just heard some rather ungrammatical English being spoken. Or was it Chinese with some English words in it? According to some linguists, it was neither. What you have just heard is a language variety known as Singlish, or Colloquial Singapore English in academic circles. Singlish is a native variety of Singapore whose lexicon and grammar in part derive from English, Chinese and Malay. It is spoken by a majority of the young generation of Singaporeans, who relish it as their native tongue. It is officially labelled not 'good English' by the government and active steps have been taken in order to discourage its use. Educators and politicians in Singapore speak negatively of it, yet use it in the classroom and in political rallies. Ask five different language experts (or linguists) what they make of such a language and you are likely to get four different answers. This is the world of contact language formation, i.e. the evolution of new languages in multilingual environments. This book deals with the many issues – social, historical, and grammatical – that such languages raise.

1.1 Introduction to contact language formation (CLF)

Why do languages change over time? What leads a community of speakers to adopt a new language or, even more interestingly, to develop a new one?

9780521863971c01_p1-17.indd 1 8/11/2009 12:00:12 PM

Language change – in particular contact-induced change – has developed into one of the most fascinating areas of linguistics since Weinreich's (1953) seminal work. There are several reasons to justify the appeal of the field: perhaps the most striking is its multidimensionality, spanning as it does traditional subdisciplines of linguistics, namely sociolinguistics, historical linguistics, structural analysis, and the ideology of language. In addition, understanding CLF, the topic of this book, holds the key to some of linguistics' most fascinating questions: How does a new grammar (and a new language) emerge? What are the socio-historical conditions that trigger it and what factors determine the grammatical outcome? What is it that speakers do in situations of language contact and why do they do it? Are there constraints on the creativity of language or is CLF completely unpredictable?

In many parts of the world, children and adults function in multilingual environments where several languages are spoken, perhaps in different domains, such as one language for the home (used predominantly in the family or one's ethnic group), one or more language(s) for the market place, and yet another language for official interethnic communication. Often these languages are mastered very early on and learnt in an informal environment, that is, without the support of school or other institutions. The strict sense of grammatical norm that speakers of Western cultures have become accustomed to in the past few centuries is not felt to the same degree in these communities, where variation between speakers along social class, ethnolinguistic group, religious sphere or generation may be very pronounced. In these multilingual environments switching between and mixing of different languages is the norm. Even in Western societies, mixing of codes can be readily observed in bilingual environments, such as within immigrant communities (think of *Gastarbeiterdeutsch* 'foreign workers' German', the variety spoken by predominantly Turkish immigrants) or in border areas (e.g. Italy's Northwest, where the Italian-German variety known as *Alto-Adige* or *Südtirol* is spoken).

In certain times and places, practices of code-mixing have eventually led to the formation of a new language. Perhaps the most familiar evidence comes from Creole languages, typical products of the displaced communities that emerged out of European colonial exploitation of, in particular, West African, Caribbean, and Pacific regions. But population movements, forced and free migrations, and population admixtures have in fact recurred in human history long before the recent European colonial era (fifteenth to twentieth centuries) and are still ubiquitous today. Indeed we can detect them in all regions of the globe for which we have good historical documentation.

The first time I experienced the fascination of a mixed language and culture was in Melaka, Malaysia. As a second-year student of Mandarin Chinese, I had already experienced the marvel of 'variation' when I encountered Cantonese in Hong Kong and later Minnan (or Hokkien) in Penang. I had expected to find

other Sinitic languages to be somewhat similar to the Mandarin I had learnt but I was instead struck by completely different grammatical systems; the Sinitic varieties were not mutually intelligible, which rendered the application of the label 'dialects' completely useless in their case. In Melaka I discovered another dimension of marvel in the Peranakan[2] culture, with its intriguing fusion of Hokkien and Malay elements, in food, rituals, language, and other cultural domains. When and how had this happened? Why had the Chinese *Babas* and the Malay-Indonesian *Nyonyas* developed this new language known as Baba Malay?[3] And how exactly could one explain the grammatical outcome? I carried these questions with me during my graduate years in Stockholm and Hong Kong, and found other, fascinating phenomena, such as Makista (or Macanese), the variety of Portuguese infused with Sinitic elements in Macau, which was, until recently, a well-preserved example of a city-port in which East and West met. The encounters that produced these mixed language varieties also produced hybrid architectural styles and innovative cultures that offer a fascinating insight into cultural and linguistic admixture.

My first academic appointment was at the National University of Singapore, where I had the opportunity to study in depth the incredible multilingual ecology of Singapore, in which Singlish and Bazaar Malay have emerged. Also, thanks to a university research grant, I was able to conduct fieldwork on Cocos Malay and Sri Lanka Malay, two interesting contact Malay varieties of the Indian Ocean. In more recent years I have been working with the Sri Lanka Malay speech community as part of Volkswagen Stiftung's initiative for the documentation of endangered languages (DoBeS). The present book reflects the knowledge that I have gained to date from this rich linguistic experience, knowledge based on case studies that I conducted on salient aspects of CLF in these Asian contexts. Some of the restructuring is still in process and was investigated first-hand in fieldwork I conducted over the years. Other varieties are no longer spoken; thus the contact scenarios had to be reconstructed through historical and typological analysis. I discuss all of them in the following pages.

1.2 Research questions

This book aims at providing a conceptual framework, as well as a number of case studies, for exploring one underlying theme: the interaction between social factors and grammatical features in the process of CLF. In order to further our understanding of how new languages emerge, it is important to study the division of labour between 'external' and 'internal' factors in the evolution of new grammars (Mufwene 2001). This means that we need to bring together observations of a socio-historical nature with grammatical (here: functional-typological) analysis, laying emphasis on the former as the main force behind

9780521863971c01_p1-17.indd 3 8/11/2009 12:00:12 PM

CLF, as advocated by Thomason and Kaufman (1988). Specifically, throughout this book, I attempt to offer some answers to the following questions:

1. What kinds of ecologies trigger the emergence of new languages?
2. What is the relationship between input languages and the new grammar?
3. How do speakers use languages in multilingual environments?
4. Why do new languages evolve?
5. Which are the theories of language that can be reconciled with the lessons learned from CLF?

Following in the footsteps of Thomason and Kaufman (1988), I seriously consider the claim that languages change because speakers change them, by looking at both speakers' agency and grammatical systems in accounting for CLF in Asian contexts. Through the study of cultural and linguistic contact in Asian contexts, I further explore the relationship between the formation of new identities and the emergence of new languages, already established by Le Page and Tabouret-Keller (1985) in the case of the West Indies. Moreover, the case studies in this book show that a unified approach to the processes of language change and language maintenance clarifies and improves our current understanding of language change, in line with an evolutionary approach to language change as exposed in Croft (2000) and Mufwene (2001). In evolutionary terms, language change and the lack of it are seen as processes that can be explained by the same cognitive mechanisms of selection and replication of utterances. Finally, this book stresses the importance of ecology, in the sense of the environment in which CLF takes place ('external ecology'), as well as the systemic interaction of the linguistic codes in contact ('internal ecology'), as already pioneered by Mufwene (2001, 2008). While at the University of Amsterdam, I have enjoyed the benefit of exchanging ideas with colleagues whose expertise lies in language contact in different regions of the world, in particular West Africa, the Caribbean, and South Africa. It is in the hard work of these (and other) colleagues before me that I find inspiration to present what I hope will be novel, useful materials to advance our understanding of the creative linguistic behaviour typical of usually multilingual diasporas around the globe.

In the approach taken in this book (see chapters 4 and 5), the underlying assumption is that CLF, including the development of pidgins and Creoles, or mixed languages in general, is a natural outcome of multilingual environments in which languages are learnt in informal settings (the home, the work place, the market), as opposed to the school. This is particularly so in the case of displaced communities – diasporas that can be regarded as 'minorities' or peripheral within the geopolitical contexts in which they are situated. I will show that in polylectal speech communities, external factors such as (a) size of the community, (b) type of intra- and intersocietal networks, and (c) multilingual practices actively influence and, for the most part, determine the outcomes of

contact situations. 'Speech community' is intended here in the sense of a web of social networks (Milroy 1987; Milroy and Milroy 1992) with group-internal relationships as well as intercommunal relationships that, to various degrees, determine the types of linguistic attitudes that define the community.

Cutting across these parameters is the typological dimension, i.e. the similarity or difference between the grammatical systems in contact ('congruence' in Mufwene 2001). This dimension is of the utmost importance as I assume that it is only by relating grammatical analysis to points (a) to (c) above that a complete picture of CLF can be achieved. In this sense, I aim to explain patterns of CLF as essentially not divorced from observations derived from historical linguistics, language variation and change, areal typology, and sociolinguistics (e.g. Siegel 2004; Heine and Kuteva 2005; Ansaldo, Matthews and Lim 2007).

The case studies presented in this book offer substantial evidence for the fact that CLF is the natural – though not necessary – outcome of population admixture in shifting socio-historical contexts. The findings present compelling evidence in support of the following claims:

- language use is inherently creative;
- multilingualism and/or casual transmission support innovative use of language;
- the typological kinship of the input languages determines to a high degree the grammatical outcome;
- CLF goes in parallel with new identity formation;
- CLF is not the consequence of a negative process (i.e. failure to acquire a target language) but rather of a constructive process involving speakers' agency; and
- CLF happens most often in minority groups, displaced groups, or otherwise geopolitically peripheral communities, who are not necessarily disempowered but often in a position of socio-cultural brokers.

1.3 The role of ecology in Asian contexts

If we acknowledge the role that ecology plays in the evolution of new varieties, it is extremely rewarding to step away from the traditional contexts in which the dynamics of CLF have been heavily researched so far. It is obvious that the lessons learned from these encounters, no matter how valuable, have limitations. In terms of both external and internal ecology, we have been looking at encounters between Western European languages and West African or Pacific languages in colonial environments within which the power relation is in favour of the Western party. This means that, in terms of linguistic typology as well as historical and social dynamics, we still have not yet explored the full range of possibilities in which CLF can occur.

In this book then, I discuss CLF in East, South and Southeast Asia, focussing on the role of Malay and Chinese in various encounters before, during, and after the European colonial era (fourteenth to twentieth centuries AD). The geographical area I cover extends from the South China Sea to the eastern half of the Indian Ocean, including the Indonesian archipelago. As has often been pointed out in traditional historical literature (e.g. Reid 2000), this region is first and foremost a climatic zone, ruled by the two monsoon seasons that determine the lifecycle of its populations in all significant aspects of life; it is also known as Monsoon Asia. Malay and Chinese varieties played an important role throughout the history of the region as the languages associated with the economically most powerful and influential territories, a status made more evident by the spice trade between the Maluku Islands and China (see chapters 3 and 4). This offers an opportunity to observe internal and external ecologies that are different from the better-known settings in which CLF has been studied, i.e. the Atlantic and Pacific regions.

Take the case of China Coast Pidgin (CCP), for example, the oldest attested Chinese–English pidgin we know of: here the linguistic encounter involves Cantonese and English. The former is a strongly isolating language; that is, it makes no use of inflections for person, number, and tense. Also importantly, its sound system is radically different from that of English, especially in its use of lexical tone. Moreover, the Chinese were by no means colonial subjects in the formative period of CCP. Quite the contrary, as I will argue, they were in a position of power, determining the local conditions of trade between them and the English. What are the social dynamics of such encounters? Was CCP the result of the Chinese failing to master English or was it the outcome of attempts by Western merchants to master Chinese? Where does the agency lie in the formation of this new language and how do we make sense of its grammar?

Ecological diversity is also conspicuous in the evolution of contact Malay varieties. For example, there were contexts of relative isolation, in which varieties such as Cocos Malay emerged that were influenced by pre-existing lingua francas, just as there were settings of intimate contact, where varieties evolved that were influenced by adstrate languages, as in the case of Baba Malay, which bears Chinese influence. The consequences of such ecological variation lie in the different grammatical outputs of the language restructuring process. Moreover, there are cases where contact spanned a very long period of time, such as that of Bazaar Malay, which was attested for roughly a millennium, and these offer us a new perspective on the nature of pidgins and lingua francas. Finally, the patterns of multilingualism observed in certain Asian ecologies are quite unlike what we know of, say, the Caribbean region, again providing us with valuable comparative material. This is, for example, the case of Sri Lanka Malay, a language formed through the practice of widespread trilingualism (see chapter 6).

9780521863971c01_p1-17.indd 6 8/11/2009 12:00:13 PM

As suggested above, using the case studies presented in the book, it will be possible to tackle the following questions in relation to CLF in Monsoon Asia:

- Does contact language formation in the Asian context present us with novel evolutionary patterns?
- What are the dominant external factors in Asian ecologies?

In order to answer these questions properly, it is important to appreciate fully the nature of the ecological factors that define the various contexts in Monsoon Asia, which are the following:

1. Slavery and manpower: the study of slavery and manpower in different regions and historical phases helps us understand the patterns of indentured labour and population movements that led to heteroglossic networks as seen in particular in the city-ports. This relates to questions of (i) mode of transmission, i.e. how new features and varieties got integrated with the already existing ones, and (ii) notions of 'target language' and feature pools available to speakers.
2. Economy and politics: seasonal trade routes determined by the monsoons gave rise to a number of multicultural and multilingual enclaves. How were the numerous languages negotiated and what were the outcomes of such negotiations in the different ports? And how did the transition of power from local networks through 'interactive' colonization (Faraclas, Walicek *et al.* 2007), involving trade, intermarriage and other forms of mutual exchanges, to exploitation colonization impact language change in the region?
3. Social relations: mixed cultures that arose in precolonial times from quasi-equal social relationships between the parties (e.g. marriages between Chinese merchants and local women in Malaysia, or communities that emerged from the mixing of Portuguese, Malays, and Chinese) provided ideal settings for the evolution of new languages.
4. Language varieties and linguistic codes: the typology of the languages involved, Malay and Sinitic, can show remarkable congruence at a superficial level. It is crucial to distinguish between written and spoken registers in both cases because there is a great divergence between the registers, and it was predominantly features from colloquial varieties that were involved in interethnic communication. It is also important to look at the appropriate varieties in each case, as there is significant regional variation. When looking at Malay influence, for example, we have to be aware of the variety of languages spoken in the Indonesian world and their significance in the past. When looking at Sinitic influences, we usually have to focus on southern varieties, Hokkien (Min) and Cantonese, typologically very divergent from Mandarin (or Standard Chinese).

In this context, I find it particularly enlightening to look at cases of CLF of precolonial times as a direct product of the social, cultural and political dynamics of Monsoon Asia. The development of pidgin Malay varieties such as Bazaar Malay (Adelaar and Prentice 1996) is a clear indication of the ubiquity of multilingualism and language admixture, and shows us that CLF can occur under historical circumstances that are not particularly extraordinary. Nonetheless, the question of why humans develop new languages over time, rather than speak like their neighbours or their ancestors, is a very intriguing one. As we will see, when social and historical conditions are ripe for a new community to emerge from the blending of previously existing ones, CLF is very likely to occur. The external conditions for CLF typically involve one or several of the following situations (see chapters 2 and 3):

(a) trade between different communities, in particular where there is a stable space where the groups converge in order to trade, i.e. a market, a port, etc.;
(b) interethnic marriage, especially when that happens in a neutral or 'third space' as in the case of the Hokkien Babas and Malay/Indonesian Nyonyas in Melaka (see chapters 3 and 7);
(c) resettlement as a consequence of migration, whether forced or spontaneous (indenture, economic, political, etc.), as observed in the Malay Diaspora of Sri Lanka (see chapter 6).

In newly formed ecologies created by such conditions, a community that has achieved a critical demographic mass and stability in the new space will often start identifying itself as a discrete cultural unit. As part of this process, there can be a societal convergence (both conscious and unconscious) towards a shared set of communicative practices that results in what appears as the evolution of a new language. For example, the case of the Peranakans of Melaka shows how a new community that arose out of intermarriages between Chinese- and Malay-speaking people came to develop a new, hybrid variety of Chinese–Malay, although both Chinese and Malay were fully available and functional in the ecology in which the Peranakans lived. The new language, Baba Malay, was not a communicative necessity or a failure in mastering either Chinese or Malay but rather an expression of the new cultural identity of the Peranakans, neither Chinese nor Malay. This is an example of why, in the remainder of this book, CLF is treated as an aspect, as well as a result, of new identity formation (chapter 7). Situations where language contact was extensive, but no new identity formation occurred, typically did not result in CLF. Extensive contact without CLF can be observed in 'linguistic areas' where the effects of contact result in typological convergence, as discussed by Gumperz and Wilson (1971) for the Indian village of Kupwar, and Ross (1996) for the case of New Guinea.

Another contribution of the case studies in this book is that, in all cases, there is substantial historical evidence to reconstruct a plausible ecology. In the cases of varieties still spoken today, such as Cocos Malay, Sri Lanka Malay and Bazaar Malay, I have been in the fortunate position of collecting first-hand data within the communities, not only relying on data elicited from informants but also recording natural language use and variation, as well as cultural data and oral history. In this sense, these case studies offer an invaluable opportunity to sketch theories based on sufficient information, rather than absence of data, as has often been the case for CLF in the Atlantic or the Mediterranean. For the varieties no longer spoken, there is often reliable socio-historical documentation for plausible reconstruction, as is the case for Baba Malay and Makista. This is discussed further in section 1.5 as well as in chapter 3.

1.4 Theory of language and CLF

In linguistics there are basically two views regarding the nature of language. In one view, our linguistic faculty is determined by innate structures that specify the range of things human language can do. In this sense, language is 'generated' by what has been identified as Universal Grammar, the biological endowment for language, or the language organ, which is common to all modern humans. In the other view, our linguistic abilities are determined by the functions of language; in this sense, language is the product of usage and, if there are universal aspects of language, they are to be found in general cognitive structures. Besides the generative view, developed by Noam Chomsky and his colleagues, and functionalist views, advanced in the work of Talmy Givón, among others, we also find intermediate views in which both positions are reconciled, as most clearly advocated in the work of Fredrick Newmeyer (see Newmeyer 2005).

Many linguists inspired by the first approach like to talk of language as an abstract entity with identifiable structural organizations that cross-linguistically have identical components. For example, they talk about 'mechanisms' of language contact and of 'acquisition' and the systemic transfer of features from one idiolect into another. However, in recent years, our increased understanding of linguistic diversity has proven time and again that universal features of language are hard to pin down, and it has been suggested that all aspects of syntactic structure may be language specific (Croft 2001: 364). For example, even the broadest, apparently universal features of language appear weak: Subject-Verb-Object (SVO) as the 'default' parameter (Kayne 1994), for instance, is severely questioned in typological literature (Dryer 1992; Hawkins 1994; see also chapter 6), and the generally assumed 'universal' noun–verb distinctions in parts of speech (Croft 2001) are questioned in particular in Gil (2000, 2001). Even more problematic is the obvious inherent diachronic change in language

that seems like a logical contradiction to the idea of a static, hard-wired linguistic 'organ' (Tomasello 2003). In fact, philosophers and psychologists, as well as evolutionary biologists, have become increasingly disenchanted by generative ideas of acquisition, and the case for innate structure is at this point at best weak, while there is an increasing amount of knowledge that argues against language-specific cognitive patterns (Tomasello 2003; Lieberman 2006).[4] Moreover, from the point of view of contact linguistics, despite attempts to identify structural constraints on contact-induced change, so far no such constraints have emerged (Koptjevskaja-Tamm to appear).

My work is very much part of the second trend and therefore incorporates a usage-based view of language and transmission; in this sense the context in which language functions is a primary domain of inquiry. The determining role of ecology in the process of CLF has been demonstrated in the evolutionary approach to language change developed in Mufwene (2001), in terms of 'internal' ecology (or the system-internal relations that can be observed when we abstract language from its social context) and 'external ecology', i.e. the socio-historical domain. When discussing the currents that affect the internal ecology of CLF, I focus on functional-typological features and cognitive aspects of language. Having established the primary role of external ecology, I assume that if any constraints are to be uncovered, they will be found in the socio-historical domain.

1.4.1 Evolution of grammar

Language evolution can be seen as the cumulation of innovations occurring in individual speech acts of individual speakers (Mufwene 2001), and languages can be seen as complex systems "in which the processes that occur in individual usage events … with high levels of repetition, not only lead to the establishment of a system within the individual, but also lead to the creation of grammar, its change and its maintenance within a speech community" (Bybee 2006: 730).

In functionalist and cognitive theories of grammar (part of the second position presented above), linguists do not isolate the structure of language from language use, as grammar is intended as the cognitive organization of a speaker's experience with language (e.g. Givón 1979a; Langacker 1987). In both cognitive and functionalist theories, grammar is seen as a set of cognitive representations that rely on general cognitive abilities of categorization, generalization, representation, etc., and are therefore not language specific. While this view has not been the dominant one in linguistic theory, in recent years a number of trends have been accumulating substantial evidence in favour of a usage-based theory of language, in particular in the field of grammaticalization studies, language acquisition, and functional-typological linguistics (Bybee

1998, 2007). As noted in Bybee (2007: 8), a very important discovery that these approaches share relates to how living languages create new grammar. Research in these areas over the past two decades shows that grammar emerges naturally out of pre-existing lexical material in language use, as is abundantly documented in the literature on grammaticalization. Grammar in this view is a constantly evolving system and "there is every reason to believe that all existing grammar came about in just the same way we observe in the documented cases at our disposal" (Bybee 1998: 250). Grammar therefore can be seen as an emergent system, complex and dynamic, similar to other complex systems observed in biology (Bybee and Hopper 2001).

1.4.2 An evolutionary take

The philosophical ideas on which my theoretical framework rests are simple: they are based on an evolutionary view of language as a dynamic, interactional phenomenon. As far as I am concerned, the evolutionary theory of language change has been convincingly presented in Croft (2000), and the ecological approach presented in Mufwene (2001) is largely compatible with it and explores further the explanatory potential of competition and selection in language change. In both approaches, which I briefly revisit in chapter 5, three common elements stand out: (1) grammars are inherently variable, i.e. even speakers of the same speech community speak differently: therefore any grammar entails variable ways of saying things (i.e. it is either multilectal or multilingual); (2) change happens through selection and non-identical replication of linguistic features, and these processes lead to feature recombinations which differ from one speaker to another, which can foster divergence in some ecologies; and (3) selection takes place in a feature pool to which all speakers of a community contribute (Mufwene 2001).

In an evolutionary approach, language change – and therefore CLF – is a result of interaction between speakers in specific ecologies. We therefore focus on the two dimensions already mentioned above:

i. External ecology: social, political, cultural, and historical patterns that may influence language use;
ii. Internal ecology: the pool of features available to speakers in the multilingual space.

Through the study of external and internal ecologies, we can determine the nature of the basic evolutionary patterns: these are selection and replication of linguistic features in the process of CLF, which are typically sensitive to factors such as prestige, focussing, systemic or statistical frequency, structural regularity, and semantic transparency, subject to the founder principle (Mufwene 1998).

Considering my emphasis on external ecology, it is natural that I investigate contact language formation at the population level (E-language) rather than at the individual level (I-language). As discussed in chapters 4 and 5, I believe that there is a methodological advantage in doing so because, if we acknowledge the fundamental role that society plays in the evolution of new grammar, it is the communal, societal output of language contact that we need to concentrate on. Moreover, I believe that in the study of past contact scenarios, we engage in a reconstruction process that can render a more informative picture at the level of E-language, or, put differently, may never be able to deliver a reliable reconstruction of an idiolect. This does not imply a firm belief in the reality of E-language vs. I-language, as I am comfortable admitting that they both are, in different ways, idealized academic constructs for the purpose of investigation. The former, however, seems to me more readily – and realistically – accessible than the latter, which has been prone to highly idealized and subjective constructions (Sampson 1997). Though it might be desirable that scholars investigating these two dimensions independently achieve similar results (Aboh and Ansaldo 2007), this may not be possible across the board. That it is useful to distinguish these domains of analysis becomes particularly clear when assessing our current understanding of the rate of contact-induced change. In E-language, change means propagation, and is clearly an external matter, which relates to type and size of society and moves gradually (chapters 6 and 8). In I-language, in a parametric view of grammar, for example, change may appear abrupt if interpreted as a matter of parameter setting (Lightfoot 1999). As shown in Mufwene (2001), this is disputable: idiolects do not change abruptly, any more than language 'acquisition' proceeds abruptly, as not all members of a class are concurrently affected by change. If we maintain a dichotomy between E- and I-language, we may thus arrive at opposite conclusions about rate of change (but see discussion in section 4.2.1 and chapter 5); in this sense it is important to note that these conclusions are not in conflict, as they pertain to different domains of enquiry, namely the study of language as a social construct vs. the study of language as an innate system.

1.4.3 Ecological factors and system-internal analysis

If we want to abstract the grammatical domain from CLF, in order to focus on the structural recombinations that occur in the interactions of idiolectal systems, it is useful to take into account the following aspects of the internal ecology:

i. Token-frequency effects: here we identify features that are prominent because they occur with very high frequency in discourse, i.e. words that are used more often or commonly.
ii. Type-frequency effects: here we look at features that are more prominent due to typological congruence between similar languages involved in the contact setting.

Both kinds of features cannot be fully understood purely at the system-internal level, even if it is there that we can isolate and study them. For example, (i) is not only a consequence of semantic-pragmatic salience but also an effect of statistical majority in a multilingual feature pool: the variety spoken by a dominant majority will be more present in discourse than others, and therefore more frequent. As for (ii), 'congruence', as defined in Mufwene (2001), indicates the degree of matching between different grammatical features; congruent features in different languages will favour maintenance in CLF, while a mismatch may undermine the survival of certain features. Congruence relates to the typological make-up of the feature pool.

In (re)constructing the internal ecology of contact, we therefore consider external factors such as numbers of speakers, their prestige, and their influence on internal aspects such as degree of congruence between systems, regularity, etc. This is why CLF is largely unpredictable, as the same relationship between comparable external and internal ecologies seldom repeats itself. All these factors can be seen at play in the formation of Sri Lanka Malay (SLM): in this ecology, the dominant language is Sinhala, both demographically and in terms of prestige, followed by Tamil, then Malay. Moreover, Sinhala and Tamil are similar to one another in many aspects of grammar, i.e. they are congruent, and this enhances both token- and type-frequencies of their features. For example, while Malay follows the Subject-Verb-Object (SVO) word order, both Sinhala and Tamil prefer the Subject-Object-Verb order (SOV). When we look at SLM, we see that the basic word order is indeed SOV. However, we also find maintenance of old Malay features in some areas of the grammar, which derive from the conflicting patterns in the adstrates that compromise selection. While in the noun phrase we see mostly Lankan grammar, in the verb phrase of SLM we find a number of Malay-like and innovative features: this can be explained by the fact that Sinhala and Tamil are far more congruent in the nominal domain than in the verbal domain. Simply put, the fact that speakers of SLM prefer verb-final word order, follow Lankan rules in marking grammatical relations, and structure events according to a combination of Malay and Lankan strategies is predictable from a close scrutiny of the social and typological aspects of all the languages involved in the multilingual ecology of Sri Lanka. In this sense, SLM is a completely 'normal' outcome of the sociolinguistic patterns and typological features that characterize the history of the SLM community.

1.5 CLF beyond exceptional evolution

In recent approaches to CLF, or language creation (DeGraff 1999), the field has become aware of a problematic tendency identified as '(Creole) exceptionalism', i.e. the tendency to offer exceptional explanations in accounting for the formation of new languages that do not fit, or at times contradict, the generalizations on language and change derived from mainstream linguistic theory

(DeGraff 2001, 2003, 2005; Ansaldo, Matthews and Lim 2007). The effects of exceptionalism are illustrated in the quote below:

> Another well-known instance of 'non-genetic' languages which did not arise as a result of a *normal and spontaneous* course of language evolution and have more than one source or parent are trade language, pidgins and creoles. Given the ongoing controversy concerning their validity as a special language type and the highly *specific social circumstances* which warrant their emergence, we choose not to expand the present volume into this field. (Aikhenvald and Dixon 2006: 11, fn.7; my emphasis)

'Exceptionalism' is the belief that there is something not 'normal' or natural about the evolution or formation of a new language. Perhaps this can be explained by the fact that Western linguistic ideology has for the longest time promoted maintenance and purism as 'normal' aspects of phylogeny; or maybe there is even a trace of postcolonial intellectual guilt to be detected here, in the exceptional nature attributed to anything arising from the conceptual space of Western colonial history. Whatever it may be, quotes such as the above justify the necessity of a volume such as this one, which aims to integrate the lessons learnt from the study of CLF with general linguistic theories. As we will see, CLF is as normal and spontaneous as any other linguistic process. And the social settings in which it occurs are by no means rare and typify much of human history, including the present time of globalization. In this sense, maintenance and innovation are not seen as 'normal' versus 'abnormal' patterns in language evolution, but rather as opposite poles on a continuum characterized by the same dynamics of language use.

The exceptional nature ascribed to contact languages and Creoles, I argue, derives from a biased understanding of the motivations behind language change (but cf. Baker 1990; Arends, Muysken and Smith 1994). Tutored acquisition in monolingual settings is quite rare in the history of the world while multilingual practices and casual transmission of language have been more prevalent by far in many of the world's societies, a recurrent theme in much of Mufwene's uniformitarian approach (2001–2008). It has been proposed that exceptionalism is partly the result of a field that has so far relied heavily on studies based on situations dominated by Western colonial powers and the ecolinguistic environments they create (DeGraff 2003; Ansaldo and Matthews 2007). In that field, Eurocentric notions that originate in nineteenth-century nation-state ideology have shaped, and still influence, our views on linguistic classification and language purism. These notions, often supported in traditional philological and historical linguistic environments, fail to capture the nature of the relationships and interaction of different languages in other parts of the world, as already pointed out in a number of studies on languages in Australia (Dixon 1997) and Laos (Enfield 2003). The limitations implicit in such a field are at least two:

1. The input–output relation. Typically, we look at encounters between isolating languages (e.g. West African) and fusional ones (Standard European) where the latter occupy a position of power. With such constant typological and social aspects, we predict similar outputs (albeit with variation) to arise from the contact environment. As mentioned in section 1.3, when we shift our enquiry to Asian ecologies, the social and typological patterns change. As we will see in chapter 2, for example, the type of 'slavery' observed in Asian contexts differs from what we know of the Caribbean region, leading to different social dynamics. Likewise, the grammatical systems in contact in this region include Chinese, Austronesian and South Asian varieties, thus offering a new perspective on the effects of typological admixture. The formation of Sri Lanka Malay, for example, shows a case of typological shift from isolating to agglutinative morphology, something that has been considered rare in CLF so far, which has led a number of linguists to formulate hypotheses relating the evolution of morphology – regarded as 'complex' structure – to time-depth (see McWhorter 2005).
2. The idealized transfer model. From a Western perspective, 'normal acquisition' involves learning in a scholastic environment, with a clear notion of standard language, and, often, very little linguistic diversity. This view has deeply influenced current theories of language transmission. But, crucially, compulsory education for the masses, standardization and monolingualism are part of what is a recent, Western European reality. They are not typical of human history, nor common across the world. The study of past, and to some extent present, Asian ecologies clearly shows that frequent and widespread multilingualism (or multilectalism) is the norm, and that casual, rather than formal, transmission is very frequent. Observing such contexts can add valuable knowledge to our understanding of patterns of second (and third) language acquisition in diverse multilingual contexts (cf. Siegel 2004). Again, the case of Sri Lanka Malay shows how typological congruence in the multilingual pool can go a long way in explaining the types of admixture that speakers undergo as a consequence of being heavily exposed to languages such as Sinhala and Tamil during the formative years of the Malay Diaspora.

The general assumption in this study is that language contact, i.e. the process of mixing of originally separate linguistic systems to create a somewhat new one, is the norm rather than the exception in the evolution of language. Social and grammatical hybridity, which I treat as the result of new identity formation, is what defines the most fundamental aspects of the continuously changing, heteroglossic communities of our world (Whinnom 1971; Croft 2001; Ansaldo 2003a). This hybridity needs to be understood in terms of degree: depending on the social and the linguistic type of the groups involved, the

output of a contact situation may resemble or depart from the input varieties to varying degrees. This is clearly illustrated by comparing different scenarios of CLF, such as Sri Lanka Malay and Cocos Malay, both varieties developed in the Malay Diaspora. In the former, in an ecology within which Sinhala and Tamil – Sri Lanka's dominant languages – are spoken, we see a process of CLF in which the socio-historical conditions trigger a restructuring process that leads to a language that looks more 'Lankan' than Malay. In the latter, an ecology within which various Malay-like languages are spoken, and where there is little socio-cultural pressure to trigger change, the grammatical output will be less divergent from the input (see chapters 6 and 7).

1.6 Outlook

This book is about the formation of new languages and new cultures and identities. Contact between humans from different ethnic and linguistic groups is one of the dominant factors for the spread of material and social cultures. Many modern cities of today rest on the remains of what used to be ancient markets, border crossings or common waterways, where practitioners of different cultures and languages came to trade. Many such spaces still exist in the non-Westernized world, and language contact as well as contact language formation are still very much part of our human landscape. It is through contact and fusion that the most creative aspects of humanity can be discerned, as the new linguistic repertoires that emerge from CLF clearly show; in this sense, contact-induced change is a very 'normal' practice for humans, and an ancient one. It is the kind of language maintenance which requires the compilation of vocabularies, elaboration of spelling and grammatical norms, schools and academic institutions and an army of fierce language purists that appears abnormal from a human communicative point of view. Therefore, in order to understand fully the nature of human language and communication, we need to first understand the natural, spontaneous and creative use found in multilingual, untutored environments, of which CLF is a central part.

Notes

1 The conversation is represented here not through standard English spelling, which would create the illusion of complete intelligibility between the language in question and English, nor through IPA, which would require linguistic training in order to be properly appreciated. In this way I hope to allow the general reader to access this short text and get the 'feel' of being exposed to a typical 'Asian English' variety. The numbers indicate pitch levels, where 5 is a high tone, and 1 is a low tone. (See chapters 5 to 7 for notes of Singlish grammar.) The conversation (courtesy of Lisa Lim) goes something like this:

-Hey, did you watch *PCK* [*Phua Chu Kang*, a local sitcom] last night?

-Yes! Wow, he was sent to English class – is his English really so bad?
-Of course! The government says one must speak good English. Don't fool around. Otherwise people will think you are really Chinese-educated.

2 This is a term used to refer to the descendants of Chinese merchants and Malay or Indonesian women born in Southeast Asia.
3 *Baba* refers to the male Peranakans, *Nyonya* to the womenfolk.
4 The innate language faculty has been recently reduced to one feature: 'recursive merge' (Hauser, Chomsky and Fitch 2002). This is a well-established feature of complex systems not unique to language (Edelman 1992).