Language Creation through Contact

Narrative Review

Author: **Pier Simone Pischedda**, Department of Linguistics, School of Languages, Cultures and Societies, University of Leeds, LS2 9JT, Leeds, UK.

E-mail: p.s.pischedda@leeds.ac.uk

Language creation through contact is an intricate phenomenon encompassing various linguistic developments such as pidgins, creoles, mixed languages, and constructed languages (conlangs). These language varieties often emerge in contexts of significant social change, such as colonisation, migration, and intermarriage, where speakers of different linguistic backgrounds need to communicate. Traditionally, linguists viewed pidgins and creoles as simplified or deficient forms of more prestigious languages, but modern research has highlighted their complexity and linguistic significance. Likewise, constructed languages have started receiving academic attention only in recent times. This report aims to provide an exploration of the diverse outcomes of language contact, examining the linguistic and social contexts that give rise to these languages, and the processes through which they evolve.

Pidgin and Creole languages

'Pidgin' and 'creole' are terms used to describe languages that emerged from language contact, often during colonisation. These varieties are of particular linguistic interest because they developed very suddenly in contact situations, where speakers with different linguistic backgrounds needed 'to fulfil basic communicative functions', such as trade (Romaine, 1988, p.2). The rapid rise and evolution of these languages have provided valuable insights into language change, acquisition, and universal grammar.

The concept of 'stratum' is frequently applied to describe languages interacting in colonisation contexts, with 'substrata' referring to the languages of the indigenous populations and 'superstrata' referring to the languages of the conquering newcomers (Filppula, 1990). Historically, pidgins and creoles were often dismissed by earlier linguists as 'aberrant' or 'defective' versions of more prestigious languages, with English-based pidgins, for example, being derogatorily termed 'broken English' (Holm, 2000, p.1). Early research defined pidgins as 'structurally deficient auxiliary languages' that eventually 'expanded into creoles' (Kouwenberg & Singler, 2012, p.284). However, this view has since been abandoned within the field, as 'each part of that definition has been challenged' (Kouwenberg & Singler, 2012, p.284). The distinctions between pidgin and creole often hinge on geographic and historical contexts. Pidgin languages are typically found in regions like the Pacific and West Africa, deriving from English, such as in Hawaii, the Solomon Islands, and Papua New Guinea. In contrast, creole languages are frequently spoken on both sides of the

Atlantic and include varieties that derive from English, French, and Portuguese, such as in Guyana, Sierra Leone, the French Caribbean Territories, and Haiti (Kouwenberg & Singler, 2012, p.284).

Despite these distinctions, there are commonalities in their development. The term ‘jargon’ or ‘pidgin’ is often used to describe the early stages of language contact, while ‘creole’ and sometimes ‘patois’ refer to languages that have undergone nativisation, becoming distinct linguistic varieties. Historically, the development of pidgins and creoles has been characterised by two main stages. The first stage, pidginisation, involves the reduction of linguistic resources and restriction of use, where speakers of different linguistic backgrounds learn a second language with limited access to the dominant group’s language (Romaine, 1988, p.2). The second stage, creolisation, involves the expansion of these dimensions, particularly when the linguistic input is received as part of first language acquisition by younger speakers (Romaine, 1988, p.2-3).

Both pidginisation and creolisation are characterised as language acquisition under restricted conditions, reflecting similar underlying phenomena and involving the development of linguistic systems in distinct contexts of language learning (Romaine, 1988, p.2). Historically, these language varieties were often perceived as dialects rather than fully-fledged languages. This perception stemmed from the fact that much of the vocabulary in pidgins came from the dominant language group, while the grammatical structure remained influenced by the native languages of the

colonised populations. The terminology used to describe these languages reflects a clear power imbalance, as early linguistic research was heavily biased, with studies typically conducted by colonisers rather than native speakers. Consequently, pidgins and creoles were frequently viewed as ‘parasitic rather than independent linguistic systems, resulting from random mixing’ (Romaine, 1988, p.7). This perspective persisted among linguists even into the 1970s, with pidgin expressions often dismissed as clumsy or amusing forms of the colonisers' language, seen as ‘debased or bastardised versions’ (Romaine, 1988, pp.7-11).

Linguists began to take notice of pidgins and creoles in the late 1800s, initially focusing on creoles from the Cabo Verde islands, referred to as Cape Verdean Crioulo (Romaine, 1988, p.4). Subsequent research expanded to include new language varieties in Africa, Asia, and America, which were collectively termed Neo-Latin. The year 1959 marked a significant milestone in the academic study of these languages with the first conference on Creole Language Studies held in Jamaica, signalling the formal recognition and scholarly interest in the field (Romaine, 1988, p.5).

Mixed languages

Mixed languages are defined as those in which different basic elements trace back to different ancestries, or as languages characterised by mixed vocabularies and morphosyntactic properties (Hickey, 2020, pp. 204-206). Scholarly opinions diverge on whether pidgins and creoles qualify as mixed languages. Pidgins typically lack

native speakers, disqualifying them from being considered mixed languages in the strict sense. In contrast, creoles do have native speakers, but their vocabulary often predominantly comes from one linguistic group, thus complicating their classification as fully mixed languages. True mixed languages are distinguished from other types of contact languages, such as creoles, by their retention of subsystems from their source languages, remaining stable and highly conventionalised ways of speaking that consistently combine elements from the source languages across different speakers, times, and contexts (O'Shannessy, 2021, p.325). These languages are typically used in highly bilingual communities and often emerge in contexts of language shift. Unlike creoles, which arise in communities where speakers are not bilingual but need to communicate with others in a different language, mixed languages involve a balanced and stable mixture of features from various linguistic varieties. Although these languages are rare, some real-world examples exist. Chindo, or Chinese Indonesian, is spoken by an ethnic group descending from Chinese male traders and local Javanese-speaking women, combining Indonesian content words with Javanese grammatical affixes (Hickey, 2020, p.213). Spanish-Philippine mestizo languages, found in northern regions of the Philippines like Vigan and Luzon, mix Spanish content words with local grammatical systems. Sri-Lankan Portuguese evolved from a creolised version of Portuguese, influenced by neighbouring languages and dialects (Hickey, 2020, pp. 212-215).

Scholars suggest two main pathways for the creation of mixed languages. One is through intermarriage, where descendants of sufficient numbers form a new ethnic

group, often adopting the mother's grammatical system and the father's lexicon (Hickey, 2020, p. 215). Another pathway is through migration, where nomads speaking one language settle in a region where another language is spoken, resulting in a language with a local grammatical system and a lexicon from their original language. Notably, not all descendants in these scenarios may end up speaking a mixed language (Hickey, 2020, p.215).

Mixed languages differ from code-switching, which involves alternating between languages within a conversation by bilingual speakers. Mixed languages require a certain historical depth and do not necessitate bilingual competence. For instance, a speaker of Chindo might not be able to speak Chinese or Indonesian separately, while code-switching would imply proficiency in both languages (Muysken, 2011, p.304).

Conlangs (constructed languages)

The world hosts a multitude of naturally occurring languages, but often overlooked are the languages that have been 'consciously created by individuals', known as constructed languages or conlangs (Goodall, 2023). These can be broadly categorised into four main types: philosophical languages, international auxiliary languages, fictional languages created for entertainment, and experimental languages used in psycholinguistic research (classification taken from: Goodall, 2023, p. 420).

Conlangs are relevant to the study of language contact because, although they are created from scratch, they often resemble naturally occurring languages in structure and are intended to be used similarly (Goodall, 2023, p.420). The field of interlinguistics studies these invented languages, which can be created either a priori, from scratch, or a posteriori, based on existing natural languages (Darquennes, 2019, p.124). Sometimes referred to as artificial languages, conlangs represent an extreme case of language contact, requiring significant conscious effort and offering clear evidence of who initially developed the language variety (Darquennes, 2019, p.124).

The earliest examples of conlangs are philosophical languages, invented in the 17th century to facilitate more precise reasoning and a truer understanding of the world. These languages were part of a broader intellectual movement aimed at creating a perfect language that could encapsulate the essence of knowledge and thought. John Wilkins' language, where the relation between a word's form and its meaning is systematic, is the most famous example. Wilkins, a prominent scholar and clergyman, sought to create a universal language that could eliminate the ambiguities and irregularities present in natural languages, thus enhancing scientific and philosophical discourse (Goodall, 2023, p.420).

In the mid-19th century, international auxiliary languages emerged, driven by the desire to facilitate communication across language barriers. The rise of globalisation and international travel underscored the need for a common linguistic ground. These languages, like Volapük and Esperanto, were designed to be easily learned with

regular morphology and vocabulary drawn from major European languages. Volapük, created by Johann Martin Schleyer, and Esperanto, developed by L. L. Zamenhof, aimed to promote international understanding and peace by providing a neutral and accessible means of communication. These became the first constructed languages acquired by a large number of people, fostering communities of speakers worldwide and inspiring subsequent conlangs (Goodall, 2023). Esperanto is seen as the most successful and widely spoken international auxiliary language. Zamenhof envisioned Esperanto as a tool to overcome the linguistic barriers that he believed contributed to national conflicts. The language's design emphasises ease of learning and neutrality, with regular and agglutinative grammar, a phonetic spelling system, and vocabulary primarily derived from Romance and Germanic languages. Esperanto's simplicity, such as its uniform verb conjugations and lack of irregularities, allows speakers to achieve proficiency relatively quickly compared to natural languages. The Esperanto community has grown steadily, with active speakers worldwide and a rich cultural life that includes literature, music, and international gatherings such as the annual World Esperanto Congress. Despite not being adopted as a universal language, Esperanto has influenced the development of other constructed languages and remains a significant example of a language created with the intention of bridging human communication divides. Today, it is estimated that there are over one million Esperanto speakers, with thousands having learned it as their first language (Goodall, 2023).

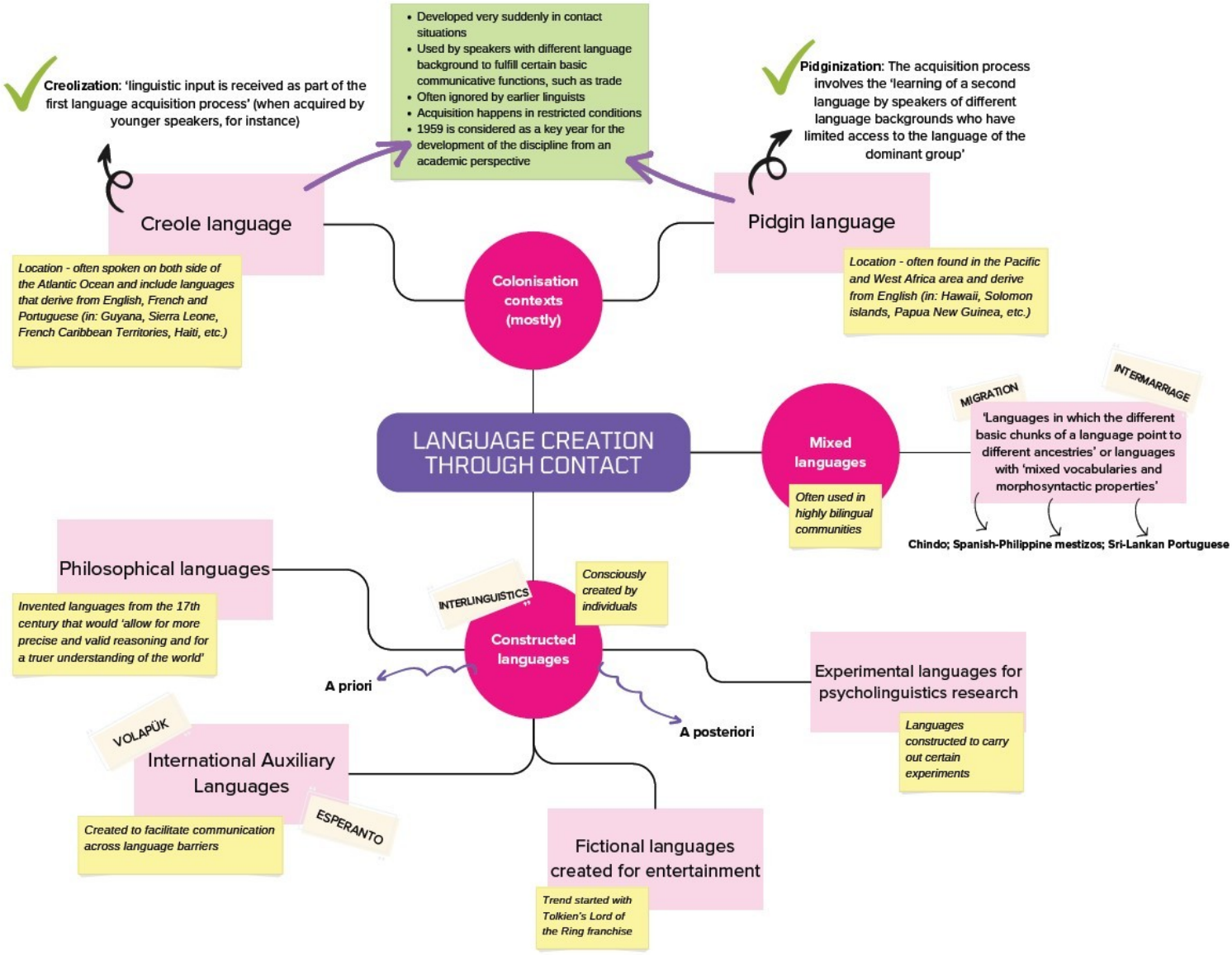
The third major phase in conlang creation came with J.R.R. Tolkien, who constructed the Elvish languages for the world of "The Lord of the Rings." Tolkien's languages, such as Quenya and Sindarin, were intricately designed with their own phonetic, grammatical, and historical systems, reflecting his deep knowledge of philology and his desire to create a richly textured fictional world. These languages not only added depth to his literary works but also inspired future creators to integrate detailed conlangs into their own stories (Goodall, 2023). Later examples include Klingon for "Star Trek," Na'vi for "Avatar," and Dothraki/High Valyrian for "Game of Thrones". Video game languages like Simlish from "The Sims" franchise further exemplify this trend. Some of these languages were initially limited but were expanded by fans, leading to increased complexity and usage. Platforms like Duolingo now offer courses in Klingon and High Valyrian, illustrating the growing interest and engagement with these fictional tongues.

Finally, experimental languages are constructed for psycholinguistic research. These languages, designed with specific characteristics, are used to study how subjects learn and use invented features. Such experiments have been conducted since the 1920s, with researchers like Leonard Bloomfield pioneering early studies. Since the 1990s, the field has expanded, utilising experimental languages to explore various aspects of language acquisition, processing, and cognitive function. These constructed languages provide controlled environments for testing linguistic theories

and understanding the mechanisms underlying human language learning and use (Goodall, 2023, p.433).

This report has examined the phenomenon of language creation through contact, focusing on pidgins, creoles, mixed languages, and constructed languages (conlangs) (see figure 1 below). By investigating the historical and social contexts that give rise to these languages, as well as the processes underlying their development, this study has highlighted the diversity and complexity inherent in language contact scenarios. The study of language creation through contact reflects the adaptability and creativity of human communication, demonstrating how communities innovate and develop new linguistic systems in response to evolving social contexts.

Figure 1. Graph illustrating various linguistic varieties emerging from language contact



REFERENCES

Adamou, E., Bullock, B. and Toribio, A. (2023). *Understanding Language Contact*. Oxford: Taylor & Francis.

Darquennes, J et al. [eds]. 2019. *Language Contact.* Volume 1. De Gruyter.

Filppula, M. 1990. *Substratum, Superstratum, and Universals in the Genesis of Hiberno-English*. Irish University Review, 20(1), 41–54. http://www.jstor.org/stable/25484338

Goodall, G. 2023. Constructed Languages. In: *Annual Review of Linguistics*. 9:419-37.

Hickey, R. (ed.) 2020. *The Handbook of Language Contact*. Cambridge: Cambridge University Press.

Holm, J. 2000. *An Introduction to Pidgins and Creoles*. Cambridge: Cambridge University Press.

Kouwenberg, S. and Singler, J. V. 2011. "Pidgins and creoles," in Mesthrie, R. (ed.) *The Cambridge Handbook of Sociolinguistics*. Cambridge: Cambridge University Press, pp. 283–300.

Muysken, P. 2011. Code-switching. In: Mesthrie, R. (ed.) *The Cambridge Handbook of Sociolinguistics*. Cambridge: Cambridge University Press, pp. 301–314.

O'Shannessy, C. 2021. Mixed Languages. In: Adamou, E. (ed.) *The Routledge handbook of language contact*. London: Routledge.

Romaine, S. 1988. *Pidgin and Creole Languages*. London and New York: Longman.

Schreyer, C. 2021. Constructed Languages. *Annual Review of Anthropology*, volume 50. Available from: https://www.annualreviews.org/content/journals/10.1146/annurev-anthro-101819-110152