Nikola Dobrić*, Martin Korenjak, Stephan Procházka,
Antonia Ruppel, Kristin Reinke, Rosemary Wildsmith-Cromarty and
Enrique Pato

Sanskrit, Classical Arabic, Latin and now English – a case of a special kind of lingua franca use and status

Abstract: This article puts forward the hypothesis that a small number of lingua francas (LFs) can be historically observed as having reached a special status and manner of use. Termed here lingua cosmopolitanas (LCs), the proposed distinctive group is exemplified by languages such as Sanskrit, Classical Arabic and Latin, representing a distinctive quasi-universal communicative phenomenon not adequately observed in research. To test this hypothesis, our study compares six LFs in terms of broad sociolinguistic characteristics. These include, on the one hand, post-4th-c. Sanskrit, post-8th-c. Classical Arabic and post-9th-c. Latin, conjectured as potentially LCs, and, on the other hand, post-18th-c. French, post-16th-c. Spanish and post-18th-c. KiSwahili in international use, as likely not of such additional status. The results of the comparison show the former as distinctive, with five characteristics separating them from the three latter LFs. They are found to include (a) loss of native speaker reference; (b) becoming the implicit conduit of the times; (c) fixedness of norm in writing; (d) extreme

---

*Corresponding author: Nikola Dobrić, Department of English, Universität Klagenfurt,
Universitätstrasse 82-67, 9020, Klagenfurt, Austria, E-mail: nikola.dobric@aau.at. https://orcid.org/0000-0002-2410-4934

Martin Korenjak, Department of Classical Philology, Universität Innsbruck, Innsbruck, Austria,
E-mail: martin.korenjak@uibk.ac.at

Stephan Procházka, Department of Oriental Studies, Universität Wien, Vienna, Austria,
E-mail: stephan.prochazka@univie.ac.at

Antonia Ruppel, Department of Religions and Philosophies, SOAS University of London, London, UK,
E-mail: antonia.ruppel@gmail.com

Kristin Reinke, Department of Languages, Linguistics, and Translation, Université Laval, Quebec City,
Canada, E-mail: kristin.reinke@lli.ulaval.ca

Rosemary Wildsmith-Cromarty, Department of African Languages and Linguistics, North-West
University, Potchefstroom, South Africa, E-mail: rosemary.comarty@nwu.ac.za

Enrique Pato, Department of World Literatures and Languages, Université de Montréal, Montreal,
Canada, E-mail: enrique.pato-maldonado@umontreal.ca

longevity in stable form; and (e) emergence of and co-existence with 'daughter' languages. The findings are then discussed in terms of paradigms describing natural language as a phenomenon in general and LF as a concept in particular. In addition, English is also discussed in the same terms, as potentially an LC in statu nascendi, and tested against the deduced signifiers.

Keywords: language status; lingua franca; lingua cosmopolitana; historical sociolinguistics

1 Introduction

Even though there is a broad body of research tackling the use of languages in international communication, one very specific kind of usage has been left largely unnoticed and unaddressed. It is exemplified by languages such as Akkadian, Ancient Greek, Sanskrit, Classical Arabic or Latin. This study argues that they exemplify different embodiments of a special status reached only by very few.

To investigate that, our study compares six languages very widely used, at different historical periods and in different locales, as international media of communication.[1] They comprise six very different languages, four from the Indo-European family, from two different branches and of various ages, one Afro-Asiatic (Semitic) language, and one Bantu language. Three are proposed as having achieved this status (Sanskrit, Classical Arabic and Latin) and three as not having reached it when in international use (French, Spanish and KiSwahili).[2] The results show that the status of the former three languages is characterised, in contrast to the latter three, by (1) loss of native speaker reference; (2) role as implicit conduit of the times; (3) very fixed norm in written material; (4) extreme longevity in stable form; and (5) emergence of and co-existence with 'daughter' languages. English is then shown to share most of these characteristics, indicating that it too may be on its way to reaching this additional status. These results are to be understood as wide-sweeping and only preliminary, principally intended to introduce and broadly substantiate a new concept and describe it in terms of a rough portrait rather than offering a conclusive empirical account of it.

---

1 Since the historical languages referred to here have been found in wide usage before the institution of nationalism, the term 'international' is utilised as an alternative to 'cosmopolitan', with the intention of implying the same meaning regarding the manner of language use.
2 Even though French and Spanish are 'daughter' languages of Latin, this does not interfere with the possibility of comparing them to Latin as they are in this study observed as LFs, i.e. as used centuries after divergence in very different contexts.

A language used internationally for communication between groups of people who speak different mother tongues is what we call a *lingua franca*[3] (LF). The term has, at present, become familiar far beyond the linguistic community generally engaged with such phenomena. What is perhaps somewhat less known outside the specialist community is that even widely-used LFs are relatively common. Dozens of them have emerged throughout known history to facilitate trade or conquest within larger or smaller geographic locales, from Persian and Mayan to Portuguese and Italian to the six languages included in this study (Ostler 2011: Ch. 2). Their social and linguistic relevance and comparative ubiquity have, unsurprisingly, prompted a large body of research, tackling LFs both from socio-economic and linguistic points of view (e.g. Crystal 1999; Firth 1996; Jenkins et al. 2011; Kachru 1992; Okombo and Muna 2017; Onysko 2016; Rose and Galloway 2019; Seidlhofer 2005; Wilton 2012; Wright 2006).

What has not been adequately covered in the existing literature is that a very small number of LFs manage to attain scopes of use spanning exceedingly large locales and wide arrays of domains, and stay in use, in truly stable form, for extraordinarily long periods of time, while most others do not (Leonhardt 2013; Mallette 2021; Wright 2004). LFs such as Akkadian, Ancient Greek, Sanskrit, Classical Chinese, Classical Arabic and Latin, as mentioned above, come readily to mind as prominent examples. The study at hand intends to focus on such LFs as potentially epitomising a special category. The hypothesis put forward here is that under certain conditions an LF can progress into a new, as of yet inadequately charted kind of language status. We term it here, for the sake of both distinction and terminological continuity, *lingua cosmopolitana* (LC).[4] This postulated language status is not observable as a function of the number of its native speakers or the political/economic power of the entity behind it (at least not in the long run). Rather, it is grounded in the dissolution of ties to native speakers, 'fixing' of the lexico-grammatical core in writing, and of an LC taking up the role of a widely shared vehicle of the times of its use (i.e. of what could be, roughly speaking, called the

---

3 The original lingua franca (from 'frank' as *Frankish* for 'European'), almost entirely unknown today outside the linguistic community was, in fact, a pidgin rather than what we understand the term to signify now. Also known as *Sabir*, it was a contact language used in the Mediterranean between the 11th and 19th centuries by sailors and traders (Brown 2022: 170). For its base it had predominantly a number of Italian dialects, as well as Spanish and Portuguese. The reason for the misnomer of what we call lingua franca today (i.e. not a pidgin but rather a third-party, 'vehicular' language), taking root as an English term, lies most likely in the perception of similarity of communicative use, namely international communication, rather than linguistic similarity of the two phenomena (Brosch 2015: 72). This is not surprising since such usage of the term appears to originally stem from a UNESCO publication which put forward that a lingua franca is any language "which is used habitually by people whose mother tongues are different in order to facilitate communication between them" (Barotchi 1994: 2211; UNESCO 1953).

4 Alternatively, 'cultural language' or 'world language' could be also used as nomenclature.

Zeitgeist) and then staying in it for exceedingly long periods of time (even beyond, or rather parallel to, the inevitable divergence into 'daughter' languages).

The challenge in tackling the newly proposed LC status lies in the fact that it, and the linguistic behaviour that an LC involves, does not correspond to our contemporary concept of a natural language, including how we conceptualise LFs. In fact, at present there is no accepted framework by which to describe an LC as a linguistic or social phenomenon. This is because an LC was last time visible (at least in the Western world, since Classical Arabic never really disappeared) several centuries ago, in the form of Latin. This absence has resulted in the non-existence of adequate academic interest in language statuses beyond those nationally referenced (e.g. first/native, second/foreign, and so on).

Both since Europe seems to be linguistically returning to its pre-modern self (in terms of how English is being used within it) and in order to address the apparent research gap, the aim of this study is to examine a number of prominent LFs in search of signifiers of LC status. This investigation is intended as the first, broad step towards a more in-depth, empirical investigation of potentially existing more concrete sociolinguistic characteristics.

2　Lingua cosmopolitana (LC)

The vast literature on LFs as vehicles of international communication (with the huge majority tackling English), agrees on several broad characteristics a language needs to exhibit in order to garner the moniker (Bamgbose 1998; Kachru 1992; Medrano 2018; Ostler 2006; Pennycook 2007; Schneider 2003; Seidlhofer 2001; Smokotin et al. 2014):

(1)　habitual use beyond the boundaries of its original community of native speakers (motivated by trade, religious, political, administrative, cultural or academic communicative needs);

(2)　emphasis on function (i.e. on its use as a means of international/intercultural communication) and not on form, which results in the appearance of many non-standard features – as a linguistic code, an LF is typically used in *appropriation* (i.e. non-native speakers employing English in ways typical to them) and with *accommodation* (i.e. with expectations of non-standard and non-native forms in it);

(3)　existence of linguistic imperialism and unresolved power-relation issues (i.e. native vs. non-native), whereby LFs fail to embody sociolinguistic neutrality despite widespread non-native use (e.g. as observable in many post-colonial settings); and, therefore,

(4)   persistence of *exonormativity*, from the perspective of international users, involving a prescriptivist focus on the speech community perceived as 'owning' the language (i.e. those we perceive as the native speakers).

In addition to this, it is also important to underline that the contemporary understanding of the phenomenon of an LF, as seen by both the academic and the non-academic stakeholders, is rooted in the *modernist* understanding of language, dominant in the last several centuries (Hogg 1996). This view of language arose alongside nationalism in Europe and has since become entrenched in sociolinguistic practice. It includes *nation*, *ethnicity* and (native) *norm* as the main reference points for any language (Modiano 1999: 25). With LFs, this is reflected in the prevailing view of, for example, English still 'belonging' to its countries and speakers of perceived primary origin in terms of nativeness, norms and standards (i.e. the UK, USA and so on).

Contrary to 'ordinary' LFs, the LC status not only implies language use beyond the boundaries of the original community of speakers. Every LC is an LF first – Latin was, for instance, perceived and used at the onset of the Roman empire in much the same way Spanish was perceived and used in the South American colonies at the height of the colonial period (Leonhardt 2013: Ch. 2). Rather, an LC comes with an added feature of extremely wide communicative spread, both geographically (in terms relative to the relevant time periods) and, more importantly, domain-wise. Firstly, languages such as Sanskrit, Classical Arabic and Latin, when in the LC stage, have been used as means of communication within vast regions and among hugely diverse peoples, especially in terms of the ancient world. For example, you could have travelled in the 12th century from Aileach Fort in Ireland all the way to the Holy Land and still managed to be understood in Latin (at least by the learned). This involved peoples who were not (anymore) connected by any kind of administrative system (such as an empire) or even a cultural or religious one. Moreover, going beyond geographic spread alone, it was accepted that one was to conduct oneself in most matters academic and literary in Latin, by default (Wright 2003: Ch. E). You could not have called yourself a person representative of your times (in the sense of contemporariness) if you were not adequately proficient in Latin.

To ascertain whether such sweeping observations (which have informed the hypothesis of the existence of LC status in the first place) can be registered when different LFs are compared and contrasted, the current study makes observations on six LFs. Three of them are historical languages considered here to have been at certain periods in history instantiations of LCs, namely Sanskrit, Classical Arabic and Latin (Leonhardt 2013: 14–15), while the three LFs considered as not having reached this status, offering a contrastive perspective, include French, Spanish and KiSwahili.

Finally, this study also engages with English as an LF considered as currently becoming an LC.

3 Methodology, analysis and results

Since this study focuses on broad similarities and/or differences between languages in international communication, related to the social context of use of a language, and since the intent is to draw on the history of their use and change as accounted for in secondary research, our approach has largely been qualitative. Following the tenets of historical sociolinguistic methodology (Hernández-Campoy and Silvestre 2012: Ch. 2, Ch. 5, Ch. 31), this involved consulting the extensive existing literature related to each of these languages and extracting a relevant description of them when in LF (or possibly LC) use. Our focus has been on searching for evidence of any salient differences between Sanskrit, Classical Arabic and Latin on one side and French, Spanish and KiSwahili on the other in terms of their behaviours as LFs. The following sections showcase the results of such an analysis, as motivated by the discussion presented in the section above.

3.1 First major distinction: loss of reference to native speakers

One of the most striking differences that can be observed between 'ordinary' LFs and the proposed LCs is that the latter had, at one point, experienced complete loss of native speaker reference in communication, i.e. the 'death' of the native speaker (Leonhardt 2013: xi). Without native speakers (that is, persons who could claim 'ownership' of a language) present in the, by definition, linguistically and culturally heterogeneous discourse community that uses an LF, belonging of an LF to the whole heterogeneous community rather than only to the 'master' (i.e. native) part of it becomes implied and accepted (Erasmus 1528; Schneider 2003). Language, in such a case, is reduced from one of (national/ethnic) *identification* to one of *communication* and convenience primarily (Hüllen 1992). The release from the ideological baggage thus allows for an easier acceptance of the language in cosmopolitan communication because this eliminates the guise of threat (linguistic and political) that is otherwise usually perceived in LFs (Motschenbacher 2016).

　　The consensus as to the 'death' of the native speaker having occurred in all three of the historical LFs analysed here is not hard to reach. After they had already been extensively used as LFs for prolonged periods of time, Sanskrit, Classical Arabic and Latin lost reference to the native speaker in both the political and linguistic (i.e. real-time communication) sense at the onset of their becoming LCs. For example, Classical

Arabic (i.e. the variety of Arabic which possesses a fully-fledged case and mood system) lost any significant reference to native speakers starting from the time of its expansion around the 8th century (Holes 2004: Ch. 1; Versteegh 2010: 61). Classical Sanskrit lost its connection to native speakers by the 4th century at the very latest (Pollock 2009: 12–14). The ultimate cause for the 'death' of the native speaker of Latin (i.e. of any persons who would claim ownership of it)[5] was most likely its church-motivated reconstruction around the 8th–9th century (Clackson and Horrocks 2010: Ch. 8). What is more, it can very well be claimed that it was this release from the bond of the native speaker reference and their implicit 'ownership' of these LFs that was crucial to their attainment of LC status. This strips an 'ordinary' LF from the threat it otherwise embodies in the perception of its non-native users as the political extension of its (usually powerful) 'owners'.[6]

Interestingly enough, reduced reference to the native speaker, even if not the 'death' itself, can also be observed with KiSwahili, but for reasons different from the loss of native speaker reference (Mugane 2015).[7] KiSwahili is actually the native language of the small ethnic group of the Swahili people on the East African coast and nearby islands in the Indian Ocean (i.e. primarily in today's Tanzania, Kenya and Mozambique). As a locally present and already habitually used LF along trading and slave routes that extended west across Lake Tanganyika and into the present-day Democratic Republic of Congo, it was co-opted in the late 18th century by the Arabic colonialism effort of the time and employed as the vehicle of communication with the different local populations. Thus, its native speakers carried no kind of power or prestige, as is usually the case at the onset of an LF. The LF also spread across a relatively wide region of Africa, far beyond the comparatively small group of its

---

5 In truth, there were faith/religious communities that made claims as to ownership of Sanskrit, Classical Arabic and Latin alike, but this was only in a weak, non-exclusive sense. In addition, those communities were superregional in the way in which a natural speech community of the modernist (i.e. nationalist) kind, a reference point to which we inherently refer today, is not (Kohn 2005: Ch. 3).
6 For example, it is not hard to imagine that a group of 15th-century European diplomats may have been happier conversing in Latin rather than in French because the use of the latter would have by default implied implicit recognition of more power being thus accorded to France, just by means of its language being selected for the international context, while for the former there was no such political baggage to consider.
7 Today, KiSwahili (or Swahili), is used as an LF throughout large parts of East Africa and the eastern Congo. KiSwahili is also one of three official languages (with English and French) of the *East African Community* (EAC) countries, as well as one of the working languages of the *African Union* and of the *Southern African Development Community*. It has official status in the Democratic Republic of Congo, in Tanzania and in Kenya, and a symbolic status in Uganda, Rwanda and Burundi (Hurskainen 2004: 364). As of recently, KiSwahili is being offered as a subject in schools, in countries including South Africa, Botswana, Namibia, Ethiopia and South Sudan. In addition, it is the educational language in Tanzania and in much of eastern Congo.

original native speakers. As a result, this special kind of use of KiSwahili left native speakers largely absent from reference from the very beginning of its use as an LF (Spear 2000).

When it comes to Spanish and French, however, the reference to the native speakers as the implicitly perceived 'owners' of the two languages has never disappeared or even diminished in perception, even at the highest points of their uses as LFs. Although French was exported to all parts of the world as a result of colonial expansion, France has continued seeing itself, and has also been perceived by others, as the single 'owner' of French (Pöll 2017: Ch. 3). This holds both linguistically (thanks to the strong awareness and policy output exercised by different French stakeholders, such as the *Académie française*) and socio-politically (thanks to the relevance of France as a geopolitical entity). Despite some questioning in the francophone world as to where the standards of French reside, they had for a long time mostly been anchored in France when it comes to its use beyond its national borders. True, since the 1960s, varieties of French outside of France have undergone processes of empowerment and legitimisation of their own linguistic standard, but only to a certain, limited extent (Pöll 2017: Ch. 4). The French language has also always been strongly linked to feelings of French nationality as well as being a core identity marker (in a positive and a negative sense) of French-speaking communities outside of France (Reutner 2017). What ended the use of French as a prestige LF (so employed between the 17th and 20th centuries in Europe) and reduced it to how it is used today as an LF (i.e. in Western and Central Africa as well as the Indian Ocean region) was not a reduction of the native speaker reference. Rather, it was the strengthening of one of its competitor LFs – namely of English – whose prestige was extraordinarily bolstered in the last century, beyond that already engendered by the British Empire, by the burgeoning power of the USA, combined with the relative loss of the geopolitical importance of France (Drake 2004; Wright 2006).

Spanish, in contrast to French, experienced a thorough adoption in most of its former colonial settings, largely the domain of its spread and use as an LF (Godenzzi 2006). An ocean apart and faced with a crumbling Spanish Empire, each former colony rose to independence around a Spanish already considered as 'owned' by many of the speakers using it in the different former colonies. This was facilitated by the fact that, rather than being self-liberated from colonial power by the subjugated population, Spanish colonies in the Americas had a tendency of having independence established by the descendants of the colonisers themselves, i.e. the so-called *criollos* (Roberts 2006). Hence, in a way, Spanish is what they felt was their native language already, despite this not being the case for much of the unprivileged part of the population (including the indigenous peoples and the descendants of slaves). As a result, rather than losing the reference to the native, Spanish developed several new native cores and was thus not really caught up in the native versus non-native

struggle observable in most other post-colonial linguistic settings (Riegelhaupt et al. 2003: 130–131). In other words, this caused Spanish not to be perceived as an imposed language. What remains of the use of Spanish as an LF is mainly in its role as the common medium within Hispanic America and in Equatorial Guinea (alongside French), as well as within the current revitalisation attempts in the Philippines (Moreno and Atilano-Tang 2023).

3.2 Second major distinction: conduit of the times

If we could glimpse a circle of learned people, most speaking different mother tongues, discussing in 9th-century Baghdad the finer points of the Qur'an, we would hear a discussion taking place in Classical Arabic. Poetry and academic work across the region would have been produced in the same language, functioning as the linguistic medium of the shared contemporariness of the Classical Arabic world of that time (Harb 2020). It is important to underline that this would have been done by default. Those scholars would not have had to discuss which language to use. There was only one prestigious expression of the 'modernity' (i.e. shared contemporariness) of their time and place, and they all spoke it as a part of their essential skillset. Making an argument as to how Latin and Sanskrit also represented the same kind of conduit of the times in their respective spheres of influence and times of employ would not be very hard (Leonhard 2013; Pollock 2009).

Looking at the three LFs proposed as likely not having reached the same kind of status and manner of use as Sanskrit, Classical Arabic and Latin, Spanish was a very commonly used language of trade, more than anything else, having been used widely for general communication in a particular time and place in relation to colonisation. KiSwahili was never used with the aura of prestige that most LFs attain, due to uneven position of power, and, hence, never for what can be called prestigious domains of communication, such as literature or academia. Only French was at one point the most preferred language of its contemporary surroundings (at the very least in Europe and North America).

Starting from the reign of Louis XIV onwards, French became the dominant LF in Europe initially, as of the 17th century, in Central Europe, and in the 18th and 19th centuries in Eastern Europe (Wright 2006). Thus, much of academia and politics involved its heavy use. Yet French has never truly achieved the position of being the only, default expression of its own times as Latin had, for example, because of its competition in Europe at the time with German and even Latin. It has also not managed to retain the prestigious role it once held (Fernández Vítores 2005: 180). Some researchers even question the widespread view of French as almost universal in the 18th century. For instance, Siouffi (2010: 411–419) points out that this perception

needs to be taken with some reserve, as it is based on a small number of texts that are often misinterpreted and inadequately confronted with reality. He highlights the fact that the spread of French beyond its borders was temporary, partial and subject to multiple dynamics, calling into question the accepted idea of its pervasiveness. Ultimately, one can say that French was potentially at the cusp of LC status in terms of this criterion, but never really there; nor is there now a chance of it attaining such a status in the foreseeable future (Drake 2004; Rjeoutski et al. 2019).

In sum, a distinctive feature of an LC is that it becomes by common understanding and unspoken agreement the preferred language of expression in any international context and domain. Using it becomes key to success and a marker of 'modernity' and professionalism. In addition, it is not only politics and academia which take place internationally in the LC by default. Perhaps even more importantly, it is the cultural output that also takes place in it. Thus, the LC becomes the intrinsic marker of the (global) community of practice of its time, representing the contemporary societies (and what is common to them) happening around it and in it. Such a community is not brought together by phenomena related to nationalism or other more traditional markers of community (such as literary tradition, ethnicity or heritage). Rather, the community of practice that finds its anchor in an LC is actually brought together by the shared 'Zeitgeist', so to speak, as the relevant 'joint enterprise', for which an LC is the *de facto* conduit.

3.3 Third major distinction: norm (very) 'fixed' in written material

The massive variability produced by immense numbers of non-native speakers (observable in glimpses in writing but very present in spoken language), prompts early on in the evolution of an LC a move towards solidifying standards (specifically of core lexicon and grammar) to combat heteroglossia. The reason for this lies in the intrinsic necessity to hold on to the utility of the language as a widely employed medium of cross-cultural communication, even if the standardisation, as in the case of Classical Arabic, overtly took the form of keeping the perceived 'holy' language pure and preserving it from 'foreign' influences and 'corruption' (Al Shlowiy 2019: 199).

What makes this process of 'fixing' an LF (Leonhardt 2013) different from standardisation observable within languages otherwise (especially today) is, firstly, that it comes across as a much more extreme process. This is because the wide international use of an LF drives language change in a fiercer and more diverse manner than any language otherwise experiences internally, even those with high numbers of non-native (i.e. migrant) speakers within (Mauranen and Ranta 2009).

For instance, if we observe any contemporary natural language, the effects of internal variation, coming in the form of sociolects, dialects, jargons and similar, are kept relatively minimal in most countries because of the internally widespread exposure to the standard language in schools and through the media. The extent of externally driven change, observable with LFs, is not confined in the same manner due to the absence of prolonged exposure to the standard, outside the foreign-language classroom. Nevertheless, *exonormativity* (Schneider 2003: 245), inherent to 'ordinary' LFs, stands as an obstacle for any non-native variability having influx into the standard.

With LCs, there are no such obstacles because of the removal of the native speaker as a reference point for the standard. This allows, and especially has done so in the past, for a free-for-all influx of non-standard forms into the LC (Dahlgren et al. 2022). As a natural reaction, intended to preserve communicative usefulness of the LC, a severe form of standardising (i.e. 'fixing') of the lexico-grammatical core occurs, orientated towards written rather than spoken material (Leonhardt 2013: 19). This may not be surprising in historical terms, where the existing literature output was the only tangible representation of language to use as reference for any standardisation effort. For example, Latin was thus 'fixed' in the literary output of the 1st c. BC/1st c. AD (Rovai 2021). Yet, even today, languages in wide international use still draw predominantly on written native-speaker language material, in the form of corpora or via AI, when defining standard language and its norms both for first language and for LF international use (Gilquin 2022). Moreover, such languages are learned in international contexts predominantly via their written representations, coming in the form of textbooks, grammars and dictionaries in formal education, despite the ever-increasing possibility of also utilising audio-visual aids in instruction and the yet-to-be measured exposure to social media or video games, facilitating *dual acquisition* of English (Rose and Galloway 2019: Ch. 6).

Fixing of the lexico-grammatical core, as has taken place with Sanskrit, Classical Arabic and Latin, was therefore a necessity to preserve their utility as LCs. The need and manner to do so was, for example, recognised by learned people of the time engaged in writing grammars and instructional material (e.g. Dionysios Thrax, Donatus or Sibawayh). It should be noted, however, that another aspect of preserving utility of an LC was the allowance of a certain amount of communicative (i.e. lexical) innovation to continuously, in an additive manner (i.e. addition taking place without the loss of redundant lexemes being incurred), occurring alongside the extreme 'fixedness' (Holes 2004: 5–6; Stotz 2002). Effectively, an LC adopted a mix between exonormativity and endonormativity as the means of survival (Schneider 2003: 249). Exonormativity remained present in the form of the written material that represented, in effect, the original standard language, as a means of assuring stable and long-term mutual intelligibility. Endonormativity became present in the guise of

additive innovation, as a mechanism assuring long-term communicative usefulness. This kind of dual approach to language maintenance did, for example, help increase the expediency of Latin as a medium of cosmopolitan communication to such an extent as to facilitate a millennium of use.

Languages such as French, Spanish and KiSwahili have not had to go so far in their standardisation efforts as to shut out any interference from spoken language and its non-standard varieties and stop them from affecting their lexico-grammatical core. The manner in which contemporary corpora are constructed and the central role they subsequently play in linguistic research, which, in turn, in a trickle-down manner ultimately informs the norm, still allow for some link between non-standard (i.e. spoken) varieties and the standard one to be maintained, even beyond the surface (i.e. lexical) level. Spanish, in this sense, exhibits standard fixtures of norming (i.e. norm subject to periodical change according to usage, as, for example, evident in the new *Ortografía de la lengua Española* [Real Academia de la Lengua Española 2010]), as would be expected of any language managed under the modernist world paradigm (Del Valle 2014). In other words, the discourses concerning the different cores of Spanish (in the form of language corpora) serve, longitudinally, as the bases for relevant dictionaries and grammars.[8] French, even though the often very strict official language policies ensure that its core elements remain very stable (Wright 2006: 39), remains in touch with the non-standard varieties via a vibrant output of the related linguistic research community (i.e. academic publications related to spoken language and variability therein). Ultimately, endonormativity comes across as more advanced in the Spanish than in the French-speaking world, where the prestige of the French standard partially slows down the process. KiSwahili has, in fact, only relatively recently undergone a more formal standardisation treatment, as it started taking on the role of vehicle for education and academia in countries using it in some official manner (Alcock and Ngorosho 2004: 5).

3.4 Fourth major distinction: extreme longevity in stable form

Given the scale of variability observable in the evolution of a language, it is a major assumption to consider almost any natural language used for millennia as one continuous linguistic entity (Auer and Voeste 2012: 282); and yet the case is different for LCs. Classical Arabic, Latin and Sanskrit have, once they established themselves in their communicatively dominant roles, carried these roles for exceedingly long periods of time, involving millennia of continuous use. In addition, not only have

---

8 Uniquely for Spanish, collated by the cooperation between the *Real Academia Española* and the *Asociación de Academias de la Lengua Española* (ASALE), comprising 24 academies around the world.

they been present for a long time, but they have existed in very stable forms. For example, unlike the intractable difficulty a speaker of modern Czech would have in following a text written in Old Czech (five or six centuries earlier), a speaker of Latin in the 15th century would have had little difficulty understanding a text produced a millennium and a half before. LCs, in their 'fixed' written forms, truly remain continuous entities in a linguistic sense, and do so for epochs.

The reasons why LCs tend to show extreme longevity in the function they perform internationally lie in the characteristics presented earlier in this paper. Firstly, the 'death' of the native speaker releases the language from political baggage, which normally keeps an LF from full adoption by its non-native speakers (Seidlhofer 2021: 396). Secondly, the fact that an LC is usually the most dominant (if not the only) expression of its times promotes investment into it (i.e. in teaching, publishing, technology, science and cultural output), which makes the LC hard to break away from (Kirkpatrick 2021: 3). Thirdly, deep-level 'fixedness' combined with surface-level pliability allows for utility across both time and space, solidifying longevity (Leonhardt 2013: Ch. 3).[9] Here, one may observe only Sanskrit, Classical Arabic and Latin as clearly showing such longevity in stable (written) form. Because longevity is so firmly anchored in 'fixedness', which Spanish, KiSwahili or even French still do not exhibit beyond regular standardisation, one cannot observe or predict longevity in such a stable form for them (either as LFs or as national languages).

3.5 Fifth major distinction: emergence of and co-existence with 'daughter' languages

The tension between the 'fixed' LC core and the ever-present and rampant variability of spoken forms has in the past always, in the very long run, resulted in *divergence*. In simple terms, divergence implies splitting of what is perceived to be one language into two or more languages (Wright 2012: 552). What happens is a change of status of co-existing varieties of one language, from being considered non-standard varieties to being seen as different languages, in sociolinguistic terms (Hinskens et al. 2009).[10]

---

9 To be sure, LCs do eventually fade out, usually in a very protracted process of replacement. It involves the slow adoption of whatever languages may be coming along offering more utility and prestige and thus being used alongside the ever-reducing usage of the LC on its way out, the fading out often taking centuries to complete (as was the case with Latin and the national languages in Europe, for example (Leonhardt 2013: Ch. 4)).

10 While pidgins and creoles do constitute a kind of 'daughter' languages, they are of different genesis than observed with LCs, emerging as results of language contact facilitated chiefly by imperialism rather than as a result of native varieties claiming 'independence'. Furthermore, unlike the gradual changes that occur through generational transmission with non-standard native

We term the descendent linguistic codes which established themselves in such a new status by means of divergence 'daughter' languages. In the past, with the three historical LCs, this happened because of both the exceedingly long periods of their use without widespread literacy to reinforce the standard variety and because of socio-political conditions appearing as conductive to divergence (Wright 2012: 562). In addition to that and in difference to similar processes of divergence having taken place with many other natural languages (including those of LF status), LCs do not break apart into the 'daughter' languages (as did, for example, Imperial Aramaic or Old Church Slavonic, fully dissolving into its many descendants) but rather continue to co-exist with them for very long periods of time.

In the pre-LC stage of its use (i.e. going up to the 8th c.), Latin has, like any other language, always had a number of well-developed non-standard native varieties existing in parallel with the 'fixed' standard (Mullen 2011). Due to education (and thus access to learning the standard, existing in written form primarily) being very limited and elitist, these non-standard varieties continued evolving almost independently of the 'fixed' standard (experiencing usual contact with the other languages spoken in each of the conquered territories) and only in spoken form (Adams 2014). Furthermore, after the spread of Latin as an LF, we also have compounding effects of 'non-native' non-standard varieties appearing, which led to *diglossia* occurring across the empire (i.e. as of the 3rd–4th c. and the *Tetrarchy*) (Leonhardt 2013: 88). This extensive variability across the Roman world (including in Europe) was eventually further complicated by several factors (Wright 1995).

Firstly, there was Alcuin of York's church-mediated 8th-c. reform of Latin, conscripted as part of Charlemagne's educational reforms (Wright 1982: Ch. 3). Secondly, prompted by said reforms and the manner in which they excluded what were until then the native speakers of Latin, there is the emergence of writing systems for these 'vulgar' varieties as a reaction (based on the Latin alphabet, though) and of the desire to use them for literary purposes. Eventually, as these varieties started taking over the functions of Latin, over a span of several centuries (14th–16th c.), and as ideas and feelings of patriotism started evolving in Europe relatively shortly after that (Hirschi 2015), they increasingly became perceived as languages themselves, resulting in what we today call the Romance (sub-)family of languages (Brown 2020; Clackson 2011). At the same time, Latin as an LC continued to exist parallel to them, unobtrusive for centuries.

---

varieties, creole languages emerge through a relatively rapid process strongly driven by specific socio-historical circumstances. For example, the Spanish-based creoles of *Chabacano*, *Palenquero*, or *Papiamento* emerged in a relatively short period of time as results of peculiar language contact conditions and began their life as pidgins, not as non-standard native Spanish varieties.

It is a fact that all six LFs analysed in this study can demonstrate effects of non-native-driven change stemming from their international (i.e. non-standard non-native) use (D'Anna and Benkato 2024; Dobrić 2024; Korenjak 2018; Leitner and Procházka 2025; Lwangale and Simiyu 2020; Ramírez-Cruz 2020; Ruppel 2017; van Putten 2020; Wetzel et al. 2023). However, the 'death' of the native speaker and the centripetal forces involved in the interplay between extreme variability in an international sense and extreme 'fixedness' (in the past also linked with elitism) has led to divergence and then co-existence only with the three considered as having also been LCs.[11] For example, we have, as noted, Latin and the Romance languages, then Classical Arabic and the modern Arabic languages (traditionally still referred to as 'dialects' (Procházka 2021: 219–220; Retsö 2013: 436; Watson 2002: 7–9)),[12] and also Sanskrit with Hindi and other modern Indian languages (Pollock 2009: Ch. 5).[13]

Due to staying exonormative and still firmly rooted in native speaker reference, and due to the time spans involved in their uses as LFs (being much shorter than with the historical LCs), neither French nor Spanish (nor KiSwahili, even though it is somewhat different in nature and behaviour (Amidu 1995)) have as LFs 'given birth' to 'daughter' languages (nor, by default, have had a chance to then co-exist with any descendants). As far as French varieties are concerned, there have been very few attempts defending the idea that they should constitute distinct languages, in a modernist sense. With Spanish, what we can see is an implicit agreement of 'unity in diversity', even if each country and region have their own particular characteristics. This is all despite some of them, like French, being very widely used as a prestige LF for several centuries.

3.6 Revisiting the assumptions

The main thrust of this study has been in proposing how certain LFs, under specific conditions, can attain a special status as an LC. This status was hypothesised to

---

11 It is important to note here that the 'daughter' languages did not emerge from the standard form of the three LCs but rather from the non-standard, spoken (i.e. vulgar) varieties existing in parallel.
12 With the exception of Maltese which origins in North African spoken Arabic varieties; however, being spoken in a completely non-Muslim context, it could gain the status of a fully-fledged language which is written in the Latin alphabet and has even become one of the official languages of the European Union (Borg and Azzopardi-Alexander 1996).
13 Some of the historical LCs which have continued being in use up to early modern times, such as perhaps Greek, may appear to not have 'birthed' daughter languages. However, the fact that we only have modern Greek today as the descendant, does not mean there were none in existence in the past. Even though they have disappeared in the meantime, varieties such as *Tsakonian, Pontic Greek, Cappadocian Greek* and *Italian Greek* had the potential of solidifying into individual languages had it not been for historical developments which put an end to their full divergence (Horrocks 2020: Ch. 1).

comprise a distinctive use of a language in international communication (i.e. distinctive from that of an 'ordinary' LF). Six LFs, used widely at different time periods, in different locales, and of very different linguistic, socio-cultural and political backgrounds were then analysed in terms of key properties in select cases. Three were initially assumed as being LCs at one point (Sanskrit, Classical Arabic and Latin) and three as likely not having reached that status (French, Spanish and KiSwahili). Table 1 summarises the results of the analysis in a simplified manner and should be read in the light of the entire discussion presented so far.

At the onset, it is interesting to note that the characteristics which researchers generally attribute to LFs in statuses of use termed here LCs have not, in fact, been flagged as relevant (Leonhardt 2013). These include the importance contributed to the power of the entity from which the LC originated, forceful imposition of the LC on the subjugated peoples, linguistic peculiarity of the LC somehow specially favouring its initial adoption for cosmopolitan use, and the numbers of its original native speakers.

Firstly, the political entities which fostered the languages which were to attain LC status generally did not force their language upon their conquered subjects. In fact, in some cases they actively discouraged use of it, Classical Arabic being a case in point (with *capitation* tax (Alawi 1966) reasons potentially being behind that). The truth is that the Roman Empire did not need to thrust Latin upon anyone because the majority of new speakers of it (especially native speakers of 'smaller' languages) wanted to adopt it (Leonhardt 2013: 45). The reason for this was not any kind of special linguistic quality of clarity, precision or range which would have made Latin originally more appropriate for cosmopolitan communication than any other natural language (although a millennium and a half of extensive use as an LC may have made it so through evolution). Rather, the reason behind their readiness to adopt Latin (and other faucets of the identity it represented) was that the use of Latin at the time signified prosperity, advancedness, and belonging to the community that it epitomised (Garnsey et al. 2015: Ch. 12), in addition to being practically useful. Secondly, neither of the recorded historical LCs disappeared after the political entity that had spawned them and sent them on their way as LFs disappeared. It can even be said that they blossomed after that. The 'golden age' of Latin, in terms of geographical spread, written output and intellectual dynamism, was between the 9th and 17th centuries, long after the old Roman Empire, of which it was a native language early on, had dissolved (Leonhardt 2013: 2). The *Amarna letters*, a collection of 382 clay tablets showing the very interesting correspondence between the Egyptian administration and the neighbouring kingdoms between roughly 1360 and 1332 BC, are written in Akkadian (one more language likely in use as an LC, in the ancient Near East), almost a millennium after the high point of the Akkadian empire (2334–2154 BC) (Moran 2002).

Table 1: Summary of all five signifiers observed across all six LFs.

Lingua cosmopolitana (LC) signifier	Sanskrit	C. Arabic	Latin	French	Spanish	KiSwahili
Loss of native speaker reference	Yes	Yes	Yes	No	No	Yes
Conduit of the times	Yes	Yes	Yes	No?	No	No
Norm (very) fixed in writing	Yes	Yes	Yes	No?	No	No
Extreme longevity in stable form	Yes	Yes	Yes	No	No	No
Emergence of and co-existence with 'daughter' languages	Yes	Yes	Yes	No	No	No
Total	5 'yes' 0 'no'	5 'yes' 0 'no'	5 'yes' 0 'no'	0 'yes' 3/5 'no'	0 'yes' 5 'no'	1 'yes' 4 'no'
LINGUA COSMOPOLITANA (LC)	**Yes**	**Yes**	**Yes**	**No**	**No**	**No**

At this point it may be opportune to raise the issue of the potential interdependency of the identified LC signifiers. For example, it may be interesting to consider whether 'fixing' of LCs came about as a consequence of the loss of the native speaker. Yet this is unlikely, as even if native speakers were present and claiming ownership in, say, 14th c. Europe, their linguistic presence in cosmopolitan communication taking place in Latin would have been, if nothing else, geographically rather limited. It would resemble the presence of, for instance, English native speakers in what is mostly non-native to non-native communication taking place in it internationally. Moreover, the historical LCs would not even have had the benefit of the continuing existence of the 'spoken' media (such as YouTube and similar) which still somewhat perpetuates the presence of native speakers of English in its global employ. The 'fixing' of, for instance, Classical Latin took place between approximately 100 BC and 100 AD, when the language of the major prose writers of the classical period of Latin literature (e.g. Cicero) became the standard. At the same time, native speakers of Latin existed until around 600 AD, using a range of non-standard spoken varieties. Rather, it is more likely that its 'fixing' came about as a consequence of the natural need to have a standard of some sort, both for safeguarding the communicative utility of an LC (i.e. as resistance to language change) and also for educational and literary purposes. In the case of Classical Arabic, whose 'fixing' took place during the 8th and 9th centuries,[14] additional motivation could be found in a perceived need to 'protect' the language of the Holy Quran, this time not only from internal language change but also from 'foreign influences'.

Similarly, we could consider the relationship between 'fixing' and divergence (i.e. 'daughter' languages) in the same manner. Interdependency between the two may be seen as implied due to the fact that the situation of having a high 'fixed' version by default created a diglossic situation early on in the development of the LCs, long before they reached this status. Nevertheless, it was only centuries later that the feelings of patriotism and, ultimately, nationalism came to be connected with the non-standard varieties existing in parallel with the LC 'fixed' version. It was at that point that the ultimate divergence and establishment of 'daughter' languages actually occurred.

---

14 In addition, although the actual time of the evolution of the so-called 'Old-Arabic language type' into the spoken varieties of Arabic is still the subject of some scholarly disagreement, it can be taken for granted that the vast majority of Arabic speakers had by the point of dissolution already lost the case and mood marking of Classical Arabic in their everyday language. Being the key characteristics of Classical Arabic and its grammatical descriptions, the 'fixing' of the rules of Classical Arabic may also have come about as a reaction to the native speakers' diminishing ability to obey the increasingly fixed rules (which, in turn, was also due, to a certain degree, to the more frequent language contact with increasing numbers of non-native speakers of Arabic).

Generally speaking, there is little to no interdependency to be observed between the five LC markers, implying an overall lack of a potential temporal order. While concatenation may exist in purely historical terms (with certain signifiers observable as taking place before the others), it does not seem to exist in terms of cause and effect. As the two previous paragraphs argue, it is likely that specific external, socio-cultural factors operating in specific time periods (such as economic advantage, prestige, patriotism and similar), in combination with certain demands coming from the communicative needs of the communities using an LC (such as need for 'fixed-ness' of a certain extent), served as independent impetuses, at different points in time, for the largely autonomous development of each of the LC signifiers listed in Table 1.

In sum, looking at Table 1 and the significance of the five criteria for describing the key aspects of the nature of Sanskrit, Classical Arabic and Latin and not of French, Spanish or KiSwahili, it can be said that the former three do seem to indicate an existence of an idiosyncratic group of LFs. The five identified characteristics appear to confirm the initial assumptions this study made of there being a special status of LC and of only some LFs reaching it. In addition, these five characteristics include both a consideration of the socio-political reasons behind one LF attaining LC status and of the sociolinguistic consequences of that. However, before the significance of these findings can be discussed in terms of what they may mean for the field of study of LFs and before the limitations and the fundamental nature of the present analysis can be underlined, one particular case intentionally omitted from the analysis above needs to be visited as currently central in the field.

4 English as a lingua cosmopolitana

English as an LF has intentionally been left out of the analysis conducted above because, as the section below will illustrate, there are grounds to consider English as an example of the return to LC status in wider international communication after several centuries of absence. The grounds for such a claim lie firstly in the fact that English is by far the most commonly used LF in the world today (Patel et al. 2023: 160). Additionally, it is used by almost two billion non-native speakers globally (far outnumbering those recognised as native speakers) to largely communicate with other non-native speakers (Patel et al. 2023: 161). In this sense, the global reach of English is a result of *spread* and not of *distribution* (whereby distribution, with established norms moving unchanged into different settings, implies intentionality, adoption and conformity, while spread implies unintentionality, adaptation and non-conformity) (Widdowson 1994: 379) and can be viewed as both being *hegemonic* and *empowering* (Saarinen 2014: 131).

Secondly, the special nature of English is, in linguistic terms, visible in the understanding of internationally used English as a performative phenomenon (Pennycook 2007: 73), a result of *languaging* activity by *linguistically multi-competent* (Cook 2016) non-native speakers. It is perceived as having a much lower degree of structural sedimentation (i.e. used with many more non-standard forms and with more variability) when used internationally than when used as a native language (Møller and Jørgensen 2009). Its international users, stuck between acting locally and communicating globally (Preisler 1999), cannot really be called foreign, native, second or non-native speakers. All these terms imply 'ownership' of language (Seidlhofer 2004) and are not applicable to internationally used English anymore. This leaves English as not belonging to any well-defined community, at least when it comes to traditional community markers (James 2008). Since much of this description appears reminiscent of what was implied earlier about LCs, it appears worth testing how English as an LF ranks against the characteristics deduced as relevant for Sanskrit, Classical Arabic and Latin as LCs.

4.1 English and the distance of the native speaker

Starting with the loss of reference to native speakers, it needs to be underlined that there are many international speakers of English that have never spoken to a native speaker of English nor visited any anglophone country (Choi 2015). This is both because of the geographic limitations present in many cases and also because the motivation behind learning English for most non-native speakers is actually to communicate with other non-native speakers (which, in fact, they mostly end up doing), of which there are very many more than native ones (Seidlhofer 2021).

A typical Vietnamese speaker of English as an LF would, for instance, have learned English from a non-native teacher of English, who has also learned it from the same kind of teacher. The teacher at hand would likely never have been to a country where English is native (or at least not beyond minimally, in a tourist sense) and would have had no or at best little direct interaction with those who we consider to be native speakers of English (Leonard 2018; Llurda 2004). The main vehicle of instruction would for both be written language material, such as textbooks and grammars, despite the moves to incorporate more audio-visual sources (Rose and Galloway 2019: Ch. 6). These examples illustrate well the typical physical distance of the native speaker of English in most international interactions involving English. Even indirectly, via multimedia the presence of the native speaker is questionable. The Vietnamese speaker of English as an LF would experience only haphazard exposure (i.e. in the form of music, film, or some predesigned instructional material),

often constituting not only inadequately comprehensive but also even less than understandable input (Song and Iverson 2018).

4.2 English and the conduit of the times

The extent to which English is used today as the vehicle of most of what we perceive as contemporary (i.e. modern) society is potentially unsurpassed by any LC preceding it, due to the technological and socio-political advantages it has experienced (Ostler 2011). Bolstered by globalisation, to such an extent as to become the linguistic anchor of the global "community of practice" (Wenger 1998), implying community markers beyond the traditional ones of nationality or ethnicity, and given wings by the Internet, English has reached such a default level of ubiquity in nearly every internationally communicated domain of human endeavour to have become almost mundane (Bolton et al. 2024).[15]

For example, in terms of the account of the LCs outlined above, we can also see, similar to the example of the 9th-century Baghdad scholars, that when academics find themselves at an international conference, they will naturally and automatically communicate in English (Bolton and Bacon-Shone 2020). This is also true when a group of school children meet at an international sporting event, or when an Argentinian tourist finds herself in a remote Greek village in need of directions.

This is similar to the earlier argument for Sanskrit, Classical Arabic and Latin. The same was not true of French, Spanish or KiSwahili, even at the highest points of their uses as LFs. Not only did their reach fail to match that of LCs, as accounted for above, but they also had to contend with (often stiff) competition, the likes of each other (e.g. French and Spanish), and then of German, English, and even Latin and Classical Arabic, in different locales (Ostler 2006: Ch. 16). English is today, just as Sanskrit, Classical Arabic and Latin were at the heights of their LC uses, without any real competition when it comes to being the primarily selected medium of international communication (Patel et al. 2023).

4.3 English and 'fixing' in writing

That the norm of English and standard English as a variety[16] are becoming increasingly 'fixed' in written material, potentially more so than other natural

---

15 Surpassing the observed historical LCs, English has also become the language of trade and commerce, going beyond the usual domains of scholarship, literature and culture.
16 The differences between what we call native varieties, such as British English, American English, Australian English and so on, in terms of the standard and norm are negligible when it comes to

languages, is evident in several aspects of their behaviour on the global stage. Firstly, standard English, used as the default norm for teaching English internationally and nationally, is not naturally spoken as a native language by anyone (as is the case for most languages, in fact). Rather, it is a formalisation of the linguistic use observable in native speaker-produced corpora of largely written language (due to reasons of practicality and resources), often far removed from the spoken non-standard uses (Dobrić 2013: Ch. 2). This focus on written language is further challenging because written language sources are by default elitist, sampled largely from educated speakers who are the ones who habitually produce texts to be collected in the first place (Gilquin 2022). Nevertheless, such corpora are at the core of applied linguistic research (Dobrić 2009) and it is applied linguistic research that ultimately informs the norm and the standards at one point in time (i.e. research findings slowly trickle down from academic output into policy outputs, involving textbooks, dictionaries, testing standards and curricula).

What makes the case different for English in comparison to most other natural languages, bringing in an extra dimension of 'fixedness,' is that, firstly, English is experiencing far more pressure from non-standard forms coming from the outside (i.e. from the non-native speakers using it internationally) than any other contemporary language (Patel et al. 2023: 161). It can even be argued, considering the migrant nature of countries such as the USA, that internal pressure from non-native variability has also been, for a long time, greater than for most languages (Clyne 2003). This implies a heightened necessity for standardisation in order to preserve communicative usefulness of a shared standard variety. Secondly, standard written English has become the default language of the Internet. More than half of all web content (52.1 %) is in English (Statista 2004), surpassing ten-fold the runner up that is Spanish (with 5.2 %). On top of that, we have online communication which in any international context happens in English first and foremost, and in writing (Lamb and Arisandy 2019).

Furthermore, the written language that comprises corpora of a similar kind and with similar problems of representativeness is becoming the basis for the norm and standards adopted by text generative AI and its various natural language processing (NLP) iterations of English (such as Chat GPT), which is the language most of them are at the moment primarily designed to process (Leiter et al. 2024). How such tools are to keep in touch with English as actually spoken and not only written (even within what we consider the native discourse community alone) and whether there are any corporate interests (i.e. on the side of the AI developers and providers) to do so, remains an open question.

---

writing (remaining mostly at the level of orthography). Some differences can be seen in regard to phonology (i.e. accent) when it comes to their spoken standard varieties.

Finally, the number of people learning English primarily from textbooks, grammar books and dictionaries (based largely on the standards and sources described previously as linguistically and socio-culturally problematic) outstrips by a margin of five-to-one the number who acquire it as a first language (Patel et al. 2023: 161). The classroom focus, for almost any of the two billion non-native users, is predominantly on written material, with limited input from audio-visual sources (Perez 2022), whereby even such audio-visual sources also tend to represent standard English as much as possible.

4.4 English and the possibility of longevity in stable form and of 'daughter' languages

Since English is arguably at the start of its journey to potentially becoming an LC (or is perhaps there already, to a large extent, but with a long path yet to travel) and because it is evolving in a world drastically different to that of Sanskrit, Classical Arabic (for a large part of its history) and Latin, its longevity in a stable form, in an international sense, can only be speculated about. The 'fixing' is possibly already taking place, especially the most recent kind related to AI and to educational standards at the international level, which can give some indication as to likelihood of stability. In addition, the international investment into English observable at every step, from it being one of the most prominent school subjects worldwide (often alongside the first language and maths) to its place in the online world and everything in between, may serve as indication as to likelihood of longevity.

The fact is that without remaining stable in the future, English may lose some of its usefulness in a global communicative sense, negatively impacting its longevity as an LC. Internationally observable variability first puts English in danger of *diglossia* appearing (Vila 2021) (i.e. school-promoted standard as 'high' vs. the vernacular international *Englishes* as 'low' variants) and then, ultimately, of too early divergence (Wright 2012). The danger to its potential longevity is, however, not only linguistically conditioned but also socio-politically. To illustrate the socio-political issues undermining any kind of unburdened international use of English, one need only consider the fierce contemporary discussion of English within the EU, which is torn between the promotion of multilingualism on one side and the popularity of English on the other (Breiteneder 2009; Motschenbacher 2016; Seidlhofer 2021).

4.5 English as a lingua cosmopolitana in the early stages

If Table 2 is offered as a summary of the five signifiers of LCs being applied to English as an LF, a significant amount of pertinence (as indicated by the discussion above) can be observed.

Unlike French, Spanish and KiSwahili as its contemporaries, English as an LF shows much more overlap with Sanskrit, Classical Arabic and Latin in terms of compliance to the five LC signifiers deduced earlier. It appears that aided by globalisation (so far led largely by the English-speaking USA) and by logically taking advantage of the Internet as an unsurpassed platform for natural spread, English may have managed to attain certain key characteristics that have eluded French, Spanish, KiSwahili (and most likely all other LFs). In addition, English managed to do all of this rather fast and rather thoroughly. In the first place, the main aspect of this spread of English involves it becoming the *de facto* conduit of the times, and, probably for the first time in recorded history, a truly global one due to factors such as globalisation and digitalisation (O'Regan 2021). This is perhaps the key characteristic indicating the belonging of English as an LF into the group of LCs. It helps advance both its prestige and its usefulness, motivating vast numbers of speakers to invest in it.

Other characteristics, such as the loss of reference to native speakers and even 'fixing' in writing, can only be seen as being in progress (hence provided in Table 2 with an accompanying question mark), as natural consequences of the nature of the international use of English, and thus not yet as conclusive. Firstly, there is the lack of complete equality between the native speaker being distant (as is the case with English) and the native speaker being completely absent (i.e. 'dead'). There is a yet inadequately charted and still wildly varying extent of exposure of non-native speakers to native English in the form of YouTube, music, films and similar media

Table 2: Previously deduced characteristics of LCs observed with English.

Lingua Cosmopolitana signifier	English	
Loss of native speaker reference	Yes?	
Conduit of the times	Yes	
Norm (very) fixed in writing	Yes?	
Extreme longevity in stable form	?	
Emergence of and co-existence with 'daughter' languages	?	
Total	1-3 'yes'	2 'unknown'
LINGUA COSMOPOLITANA	**Yes?**	

that international speakers of English as an LF do experience. What is equally uncharted is how this relatively limited exposure impacts its use.

Secondly, 'fixing' in writing is also an ongoing development, still offset by the linguistic community and the way it observes spoken language and expresses applied linguistics results, i.e., the manner in which investigation into non-standard spoken language varieties trickles down into considerations of standards (Altendorf 2017; Carter and McCarthy 2017; Kerswill 2006). When it comes to English, the research done on spoken language and the rich output relevant to English language teaching still potentially exert an effect on how norms are formed. Yet to what extent this is impactful is not adequately covered by the literature. Finally, longevity, stability of form, or any divergence looming on the horizon are aspects of English as an LC that only time will reveal.

5 Consequences of the findings

If all the results are summed up once more, this time including the examination of English as the seventh LF case, several conclusions can be offered. The first is that there seems to be merit in observing certain LFs as characteristically different from most others in terms of the special status they have each held at certain periods of time, indicating the validity of the original hypothesis. This is an aspect of LFs as a phenomenon that has neither been suitably observed nor accounted for, nor is there a descriptive framework by which this could be done.

The second conclusion that can be offered is that the five characteristics put forward as strongly marking LC status have shown themselves to be distinguishing enough to indicate English as also belonging to the LC group of LFs, even if only halfway there. These two conclusions come with meaningful consequences, both to how we observe language in general (i.e. within the modernist paradigm) and how we perceive LFs, most strongly observed today through the prism of English.

5.1 Updating the paradigm of what a language can be

In relation to the current linguistic paradigm of natural language as a broad phenomenon, it is evident that LC as a concept does not fit many of the characteristics of what we call a language today, especially in the socio-political sphere (i.e. in terms of possible language statuses of first, second, foreign, or even LF). Namely, as elaborated in detail below, LCs fail to conform to any of the tenets of the currently prevailing modernist view of language, with central roles still played by native-

speakerism, pre-eminence of the nationality/ethnicity references, and the modern language standardisation approaches to LFs housed in exonormativity.

5.1.1 Lingua cosmopolitana and the modernist perception of language

An LC, in the true meaning of the term, is at first glance an expression of rather an idealistic notion, when observed from our modernist (i.e. intrinsically nationalist) point of view. It implies a communicative code that is used "throughout the entire world as a neutral, human language, not intruding upon the personal life of peoples and in no way aiming to replace existing national languages". It "would give to people of different nations the ability to understand each other and would be able to serve as a conciliatory language of public institutions." Furthermore, "the primary master of this language is the whole world." It may be no surprise that these quotes come from the *Declaration of Boulogne* (Zamenhoff 1905: 1), defining *Esperantism* as a movement. What the story of Esperanto illustrates is that a crucial aspect when discussing a language in a broader forum than an academic community is the perception of 'ownership'. This can be very well shown in the example of English.

The report published by the *League of Nations* in 1922 comments in support of Esperanto on how, as an artificial language, it would help avoid the "attempt to establish the supremacy of one national tongue over all others." (Zamenhoff 1905: 1). This question of 'belonging' is precisely the issue why English, proposed as being a contemporary LC in the making, is still, at least in perception and policy, far removed from the ideal that Esperanto strove to embody. Consider that in June 2016, a few days after the British Brexit referendum, a paradoxical debate about language took place in the EU Parliament. It addressed the question whether English should continue to be an official EU language. What made it paradoxical was that those arguing for the removal of English asserted in English (in the words of MEP Danuta Hübner) "If we don't have the UK, we don't have English" (Goulard 2016). The statement flies in the face of communicative reality and illustrates the conceptual and legal vacuum around the idea that a language can exist independently of state or nation.

The MEPs felt able to discuss English solely within the framework of 'one nation-state equals one language', even though this clearly no longer fits the mould. Even more, English was held up as a threat, a painful reminder of the UK's continued and now unwanted presence. Such an ideologically driven understanding of language has recently been strengthened by the re-emergence of strong nationalism across Europe. It is remarkable to reflect on the possibilities that thousands of international workers in EU institutions might have been forced to stop using English had a loophole not been found thanks to Ireland and Malta (in the latter English is the second official language besides Maltese). It is even more remarkable, however, that

anybody felt that any loophole was needed in the first place. This is precisely the core of the problem that a language status such as LC could address.

5.1.2 Lingua cosmopolitana and the perception of language statuses

When it comes to compliance to currently accepted language statuses (Cook 2012), we can see that LCs do not conform to the contemporary manner of description. Consider how, on a national level, citizens of what are traditionally perceived as the homelands of English (the UK, USA, Australia and so on) would call it, using contemporary terminology, their native or first language. Internationally, to policy stakeholders and most of its international users, English is a foreign or perhaps a second/additional language (Cook 2016). The overly-simplistic modernist rationale for this is clear – it comes from a specific country or countries which is/are foreign to mine.

However, observed supranationally, English is used largely in a manner not fitting either of those entrenched statuses. Firstly, the folk understanding of a native language as the language of the country in which someone lives and a foreign language as one learned in school for use abroad is entirely political, and despite being very widespread and entrenched, is also objectively false. It conflates location and function (constituting, in folk terms, *nationality*), glossing over the fact that, for example, a child born to Spanish-speaking parents living in Germany could still be native in Spanish in terms of language proficiency. Similarly, English is not exclusively used outside the countries in which it is learned in school as supposedly foreign, quite the opposite. It is the most ubiquitous expression in most countries after the first language. Moreover, an 'affective' appropriation of an additional language (what English, considering the manner and extent to which it is globally learned and used today (Patel et al. 2023), likely is by default for most people today) can lead it to function within the communicative world of the speaker (i.e. learner) and its thinking at a level so close to that of the first language (even to replace it completely, partially or occasionally) that it ends up becoming a part of the identity characterisation of the individual (Hogg and Reid 2006).

Furthermore, 'first' is usually implicitly understood as being better than 'second', the latter being perceived as somehow coming after the former (Cook 2003: 3), not only in the acquisition order implied by the terminology, but also according to the value generally assigned to them. Finally, it is becoming impossible to claim that the characteristic and systematic uses of English by two billion international speakers should be dismissed as all mere errors, and all of them as mere learners of what the original native speakers somehow unalienably possess and have the sole right to prescribe (Jenkins et al. 2011).

5.1.3 Lingua cosmopolitana and primacy of spoken language

In addition to the current palette of language statuses not being applicable to how English is used internationally today, the current focus of linguistic inquiry related to it still being primarily (with some exceptions, e.g. Mauranen (2021) or Laitinen (2018)) placed on spoken language as the most legitimate object of linguistic analysis (in a sociolinguistic sense) also needs to be re-examined.

Namely, contemporary sociolinguistics operates on the belief that because synchronic written language is, in most contexts, produced (and often subsequently edited) almost exclusively in the form of standard language, it is the spoken language that is more representative and best revealing of the existing variation, allowing for more valid observation of sociolinguistics variables (Wardhaugh and Fuller 2021: Ch1). While this may be true for most natural languages existing within the present-day language policy frameworks, the case of English as an LC, with not only its immense presence in the online sphere of communication but also with its growing incarnation as an expression of NLP solutions and, ultimately, text generative AI, may imply a need for broadening our perspective.

Firstly, unlike natural languages used within the discourse confines of native-speaking settings, English as an LC leaves much more palpable evidence of its non-standard nature in writing, in its online instantiations. Written language of native speakers, even when used in informal settings, bears the hallmarks of standard English much more than that of non-native speakers (who largely lose touch with standard English upon completing schooling). Hence, written LC English may provide for an excellent insight into non-standard non-native use of it, in difference to the limited information in this respect being provided by native-speaker written language.

Secondly, the ever-growing importance and use of text generative AI (such as Chat GPT) and machine translation, both in native and non-native settings, implies that it is the language produced by such NLP solutions which will be implicitly shaping the norm of English in the future. This means that rather than having applied linguistic work stemming from native-speaker corpora inform the carriers of standard English (i.e. textbooks and dictionaries) as till now, it will be the written language produced by AI which may end up dictating the manner in which English will be used in the future (especially online, internationally).

Put more succinctly, the consideration of spoken language being much more sociolinguistically revealing holds true for considering English language use in the majority of possible statuses (Kortmann and Lunkenheimer 2013). Nevertheless, when observed as an LC, we should recognise that written language produced by all types of English language speakers (i.e. native, non-native, and, ultimately, AI), may also be, in a conceptual sense, adequately revealing of its many of its core aspects.

Such a realisation could open a path towards a renewed view of adequate socio-linguistic validity of written language (especially that representing an LC), along the lines of how diachronic written language has always been approached by historical sociolinguistics.

5.2 Updating the paradigm of what an LF can be

Most of the current research into the phenomenon of an LF is performed by tackling English as the most prominent contemporary embodiment. Likewise, most currently available research on LF paradigms is intrinsically biased by the ingrained modernist view of language discussed above. This is particularly true when it comes to questions of status. While descriptive accounts of English language forms and linguistic change as an LF can be considered ideologically unbiased (e.g. Rose and Galloway 2019), their interpretation in light of who contributes to the form the language takes, of how they do that, and of the perceptions of the relationship between proficiency and 'nativeness' is not. This is not surprising, considering that language is a phenomenon fully shaped by its interaction with its own socio-cultural environment (Haugen 1972), which makes the perceptions of the general (i.e. non-academic) stakeholders more impactful than the actual realities of language use.

For instance, the extensive research done on the use of English as an LF, both in terms of linguistic characteristics and in terms of stakeholder attitudes, has still not managed to lead us to a definite account of how English used internationally is to be observed, neither within the linguistic community (despite a number of more recent promising accounts, e.g. Jenkins and Mauranen (2019); Kuteeva (2023); or Kuteeva et al. (2020)) nor among the more general stakeholders (i.e. international speakers and related policymakers). On the one hand, most of the linguistic community still largely perceive English as predominantly and firmly anchored in the native speakers and try to approach it from the point of view of a natural language just like any other, both within the *World Englishes* and the *English as a Lingua Franca* (ELF) paradigms (Haswell 2013).

On the other hand, policymakers are stuck in the view of 'nation equals language' and 'native equals norm'. This is despite such views having massive negative consequences, from, for example, native-like standard English being the unsuccessfully sought language acquisition aim in most international educational contexts to the prestige still accorded to native speakers in the classroom over the local non-native ones, who actually present more achievable role models (Rose et al. 2021). Finally, despite the plainly observable distinction between how omnipresent English is in comparison to any other foreign language present in most communities, for the majority of general stakeholders there is little in terms of awareness of the

conceptual otherness of English. This is despite the fact that *national multilingualism* (Stewart 1968), or rather national bilingualism with English as the default additional, *adoptive* language (Fiedler 2011; Maalouf 2008) for everyone, is the already existing unofficial model in most settings. All of these issues could potentially be addressed by considering English as of novel status (i.e. contemporarily novel, as an LC). It would highlight it as a language on its way to existing beyond a reference to nationality or nativeness, as several other languages before it. This would also imply including all of its speakers, native and non-native alike in the composition of the standards and the norm (something that was possible only on a lexical level for other LCs in premodern times).

Therefore, one possible path to successfully dealing with both English and with the phenomenon of internationally used languages in general (in terms of updating the current palette of statuses) would be to consider that they are not only internationally widely used foreign languages but perhaps something more. A way forward appears to be in looking at the times when languages similar in function to what is observable in English were the norm. Surprisingly, this was actually most of the time, whereby an absence of an LC we have seen in the last several centuries (at the very least, in Europe) is actually a historical exception (Gordin 2015). Therefore, lessons need to be drawn from historical precedents, such as Sanskrit, Classical Arabic and Latin indicated here (and more, likely including Akkadian, Ancient Greek or Classical Chinese). By observing their behaviour and how it may have differed from how we observe language today, the road to an updated definition of what a language in international communication can be and what English likely is and how it can be tackled may be found.

6 Limitations and further work

The first and most serious limitation of the study at hand is that it is housed in what are mostly qualitative findings, drawn from secondary sources. Not only does the validity of the presented testing of the hypothesis depend on the validity of the five established signifiers, it also depends on the studies drawn upon offering valid results themselves. The second limitation lies in the fact that only six LFs have been chosen for the original investigation and a wider sample may have shown different results. The third limitation is related to the historical paradox concerning the validity of comparing the uses of languages of very different temporal, spatial and cultural backgrounds, even if they arguably embody the same distinctive communicative phenomenon. Finally, the study must be understood as exploratory first and foremost, intended to present a new concept and describe it in terms of broad characteristics of LCs rather than aiming to offer a definitive empirical account of it.

The limitations lend themselves well to recommendations for further work. It is clear that all of the reasoning presented above, couched in uniformitarianism, requires rigorous empirical examination and testing against source material. Therefore, any follow-up study would need to primarily focus on investigating the existence of similarity between languages proposed as being LCs on a concrete linguistic level (i.e. grammar, lexis and discourse). Here, the functionalist assumption is that the apparent similarity of use and status should produce similarity of variability and change patterns. If languages proposed as being LCs are truly embodiments of a quasi-universal communicative phenomenon, then each of them should exhibit similar patterns of sociolinguistic evolution (e.g. similar kinds of simplifications linkable to similar types of speakers or similar lexical patterns stemming from similar registers of use) despite the historical paradox. Such an analysis would be integral to bringing about the development of new terminology, new methods, and new approaches to conceptualisation of this communicative phenomenon, no matter how apparently counterintuitive to modern humans it may be.

References

Adams, James. 2014. *Social variation and the Latin language*. Cambridge: Cambridge University Press.

Al Shlowiy, Ahmed Saad. 2019. Language, religion, and communication: The case of Islam and Arabic in the Asia-Pacific. *Journal of Asian Pacific Communication* 32(2). 198–213.

Alawi, Anwar. 1966. The "Jizyah" or the capitation tax. *Journal of the Pakistan Historical Society* 16(2). 182–189.

Alcock, Katie & Damaris Ngorosho. 2004. Interaction between phonological and grammatical processing in single word production in Kiswahili. *Language and Speech* 47(1). 1–30.

Altendorf, Ulrike. 2017. Chapter 9: Estuary English. In Alexander Bergs & Laurel Brinton (eds.), *Varieties of English*, 169–186. Boston: De Gruyter Mouton.

Amidu, Assibi Apatewon. 1995. Kiswahili, a continental language: How possible is it? (part I). *Nordic Journal of African Studies* 4(2). 21.

Auer, Anita & Anja Voeste. 2012. Grammatical variables. In Juan Manuel Hernández-Campoy & Juan Camilo Silvestre (eds.), *The handbook of historical sociolinguistics*, 253–271. Hoboken: Blackwell Publishing.

Bamgbose, Ayo. 1998. Torn between the norms: Innovations in world Englishes. *World Englishes* 17(1). 1–14.

Barotchi, Maher. 1994. Lingua franca. In Ronald Asher (ed.), *Encyclopedia of language and linguistics 4*, 2211. Oxford: Pergamon Press.

Bolton, Kingsley & John Bacon-Shone. 2020. The statistics of English across Asia. In Kingsley Bolton, Werner Botha & Andy Kirkpatrick (eds.), *The handbook of Asian Englishes*, 49–89. Hoboken: Wiley-Blackwell.

Bolton, Kingsley, Werner Botha & Benedict Lin. 2024. *The Routledge handbook of English-medium instruction in higher education*. Oxfordshire: Routledge.

Borg, Albert & Marie Azzopardi-Alexander. 1996. *Maltese*. London: Routledge.

Breiteneder, Angelika. 2009. English as a lingua franca in Europe: An empirical perspective. *World Englishes* 28. 256–269.

Brosch, Cyril. 2015. On the conceptual history of the term lingua franca. *Apples – Journal of Applied Language Studies* 9(1). 71–85.

Brown, Joshua. 2020. Language history from below: Standardization and koineization in Renaissance Italy. *Journal of Historical Sociolinguistics* 6(1). 20180017.

Brown, Joshua. 2022. On the existence of a mediterranean lingua franca and the persistence of language myths. In Karen Bennett & Angelo Cattaneo (eds.), *Language dynamics in the early modern period*, 169–189. Oxfordshire: Routledge.

Carter, Ronald & Michael McCarthy. 2017. Spoken grammar: Where are we and where are we going? *Applied Linguistics* 38(1). 1–20.

Choi, Lee Jin. 2015. Revisiting the issue of native speakerism: 'I don't want to speak like a native speaker of English. *Language and Education* 30(1). 72–85.

Clackson, James. 2011. The social dialects of Latin. In James Clackson (ed.), *A companion to the Latin language*, 503–526. Cambridge: Cambridge University Press.

Clackson, James & Geoffrey Horrocks. 2010. *The blackwell history of the Latin language*. Hoboken: Wiley-Blackwell.

Clyne, Michael. 2003. *Dynamic of language contact. English and immigrant languages*. Cambridge: Cambridge University Press.

Cook, Vivian. 2003. The changing L1 in the L2 user's mind. In Vivian Cook (ed.), *Effects of the second language on the first*, 1–18. Bristol: Multilingual Matters.

Cook, Vivian. 2012. Going beyond the native speaker in language teaching. *Tesol Quarterly* 33(2). 185–209.

Cook, Vivian. 2016. Premises of multi-competence. In Vivian Cook & Tyne Wei (eds.), *The Cambridge handbook of linguistic multi-competence*, 1–25. Cambridge: Cambridge University Press.

Crystal, David. 1999. The future of English. *English Today* 15(2). 10–20.

D'Anna, Luca & Adam Benkato. 2024. On Middle Persian interference in early Arabic prose: The case of Kalīla Wa-Dimna. *Journal of Language Contact* 16(1). 1–33.

Dahlgren, Sonja, Alek Keersmaekers & Joanne Stolk. 2022. Language contact in historical documents: The identification and co-occurrence of Egyptian transfer features in Greek documentary papyri. *Journal of Historical Socio Linguistics* 8(2). 325–356.

Del Valle, José. 2014. The politics of normativity and globalization: Which Spanish in the classroom?. *The Modern Language Journal* 98. 358–372.

Dobrić, Nikola. 2009. Corpus approach as 'the' paradigm of language research. *Philologia* 7. 31–41.

Dobrić, Nikola. 2013. *The theory and practice of corpus-based semantics*. Tübingen: Gunter Narr Verlag.

Dobrić, Nikola. 2024. Effects of errors on ratings of writing performances in rating contexts in which focus on them is overtly discouraged. *Writing Assessment* 59. 100806.

Drake, Helen. 2004. France in free fall? French perspectives on the astérix complex. *French Politics* 2(2). 221–233.

Erasmus, Desiderius. 1528. *Ciceronianus*. Basel: Johann Froben.

Fernández Vítores, David. 2005. El francés como lengua franca de la Unión Europea: luces y sombras de una estrategia fallida. *Revista de Filología Románica* 27. 179–205.

Fiedler, Sabine. 2011. English as a lingua franca – a native-culture-free code? Language of communication vs. language of identification. *Apples – Journal of Applied Language Studies* 5(3). 79–97.

Firth, Alan. 1996. The discursive accomplishment of normality: On 'lingua franca' English and conversation analysis. *Journal of Pragmatics* 26. 237–259.

Garnsey, Peter, Richard Saller, Jaś Elsner, Martin Goodman, Richard Gordon, Greg Woolf & Marguerite Hirt. 2015. *The roman empire: Economy, society and culture*, 2nd edn. Berkeley: University of California Press.

Gilquin, Gaëtanelle. 2022. One norm to rule them all? Corpus-derived norms in learner corpus research and foreign language teaching. *Language Teaching* 55(1). 87–99.

Godenzzi, Juan. 2006. Spanish as a lingua franca. *Annual Review of Applied Linguistics* 26. 100–124.

Gordin, Michael. 2015. *Scientific babel: How science was done before and after global English*. Chicago: University of Chicago Pres.

Goulard, Hortense. 2016. English will not be an official EU language after brexit, says senior MEP. *Politico*. https://www.politico.eu/article/english-will-not-be-an-official-eu-language-after-brexit-senior-mep/ (accessed 04 05 2024).

Harb, Lara. 2020. Arabic literary theory. *Oxford research encyclopedia of literature*. https://doi.org/10.1093/acrefore/9780190201098.013.989.

Haswell, Christopher. 2013. A global model of English: How new modelling can improve the appreciation of English usage in the Asia Pacific region. *Asia Pacific World* 4(2). 122–137.

Haugen, Einar. 1972. The ecology of language: Essays by einar haugen. In Anwar S. Dil (ed.), *Selected and introduced*. Stanford: Stanford University Press.

Hernández-Campoy, Juan Manuel & Juan Camilo Silvestre (eds.). 2012. *The handbook of historical sociolinguistics*. Hoboken: Blackwell Publishing.

Hinskens, Frans, Kallen Jeffrey & Johan Taeldeman. 2009. Merging and drifting apart. Convergence and divergence of dialects across political borders. *International Journal of the Sociology of Language* 2000(145). 1–28.

Hirschi, Caspar. 2015. *Wettkampf der Nationen*. Göttingen: Wallstein Verlag.

Hogg, Michael. 1996. Identity, cognition, and language in intergroup context. *Journal of Language and Social Psychology* 15(3). 372–384.

Hogg, Michael & Scott Reid. 2006. Social identity, self-categorization, and the communication of group norms. *Communication Theory* 16(1). 7–30.

Holes, Clive. 2004. *Modern Arabic: Structures, functions and varieties. Revised edition. Georgetown classics in Arabic language and linguistics*. Washington D.C.: Georgetown University Press.

Horrocks, Geoffrey. 2020. *Greek: A history of the language and its speakers*. Hoboken: Wiley-Blackwell.

Hüllen, Werner. 1992. Identifikationssprachen und Kommunikationssprachen: Über Probleme der Mehrsprachigkeit. *Zeitschrift für Germanistische Linguistik* 20(3). 298–317.

Hurskainen, Arvi. 2004. Loan words in KiSwahili. In Katrin Bromber & Birgit Smieja (eds.), *Globalisation and African languages: Risks and benefits*, 199–218. Berlin: De Gruyter Mouton.

James, Allan. 2008. New Englishes as post-geographic Englishes in lingua franca use: Genre, interdiscursivity and late modernity. *European Journal of English Studies* 12(1). 97–112.

Jenkins, Jennifer & Anna Mauranen. 2019. *Linguistic diversity on the EMI campus: Insider accounts of the use of English and other languages in universities within Asia, Australasia, and Europe*. Oxfordshire: Routledge.

Jenkins, Jennifer, Alessia Cogo & Martin Dewey. 2011. Review of developments in research into English as a lingua franca. *Language Teaching* 44(3). 281–315.

Kachru, Braj. 1992. World Englishes: Approaches, issues and resources. *Language Teaching* 25(1). 1–14.

Kerswill, Paul. 2006. RP, standard English and the standard/non-standard relationship. In David Britain (ed.), *Language in the British isles*, 1–18. Cambridge: Cambridge University Press.

Kirkpatrick, Andy. 2021. *The Routledge handbook of world englishes*. Oxfordshire: Routledge.

Kohn, Hans. 2005. *The idea of nationalism – a study in its origins and background*. New Brunswick: Taylor and Francis.

Korenjak, Martin. 2018. Latein und die europäischen Volkssprachen in der frühen Neuzeit. *Philologia Classica* 13. 99–111.

Kortmann, Bernd & Kerstin Lunkenheimer. 2013. *The electronic world atlas of varieties of English*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Kuteeva, Maria. 2023. *Tension-filled English at the multilingual university: A Bakhtinian perspective*. Bristol, Blue Ridge Summit: Multilingual Matters.

Kuteeva, Maria, Kathrin Kaufhold & Niina Hynninen. 2020. Introduction: Language perceptions and practices in multilingual universities. In Maria Kuteeva, Kathrin Kaufhold & Niina Hynninen (eds.), *Language perceptions and practices in multilingual universities*, 1–24. London: Palgrave Macmillan.

Laitinen, Mikko. 2018. Placing ELF among the varieties of English: Observations from typological profiling. In Sandra Deshors (ed.), *Modelling world Englishes in the 21st century: Assessing the interplay of emancipation and globalization of ESL varieties*, 109–131. Amsterdam: John Benjamins.

Lamb, Martin & Fauziah Arisandy. 2019. The impact of online use of English on motivation to learn. *Computer Assisted Language Learning* 33(1–2). 85–108.

Leiter, Christoph, Lan Zhang, Yanran Chen, Jonas Belouadi, Daniil Larionov, Vivian Fresen & Steffen Eger. 2024. ChatGPT: A meta-analysis after 2.5 months. *Machine Learning with Applications* 16. 100541.

Leitner, Bettina & Stephan Procházka. 2025 (forthcoming). Linguistic dynamics of contact-induced change in Arabic. In Jeroen Darquennes, Joe Salmons & Wim Vandenbussche (eds.) *Language contact: An international handbook*, Vol. 2. Berlin & Boston: De Gruyter.

Leonard, Josie. 2018. Beyond '(non) native-speakerism': Being or becoming a native-speaker teacher of English. *Applied Linguistics Review* 10(4). 677–703.

Leonhardt, Jürgen. 2013. *Latin: Story of a world language*. Cambridge, Massachusetts: Belknap Press.

Llurda, Enric. 2004. Non-native-speaker teachers and English as an international language. *International Journal of Applied Linguistics* 14. 314–323.

Lwangale, David Wafula & Sarah Simiyu. 2020. Effect of mother tongue on learning Kiswahili language in public secondary schools in Kwanza sub-county, Trans-Nzoia county, Kenya. *IJRDO – Journal of Educational Research* 5(2). 63–75.

Maalouf, Amin. 2008. *A rewarding challenge. How the multiplicity of languages could strengthen Europe*. Brussels: European Commission.

Mallette, Karla. 2021. *Lives of the great languages – Arabic and Latin in the medieval Mediterranean*. Chicago: Chicago University Press.

Mauranen, Anna. 2021. Pooling perspectives: Introduction. In Anna Mauranen & Svetlana Vetchinnikova (eds.), *Language change: The impact of English as a lingua franca*, 1–11. Cambridge: Cambridge University Press.

Mauranen, Anna & Elina Ranta. 2009. *English as a lingua franca: Studies and findings*. Newcastle: Cambridge Scholars Publishing.

Medrano, Juan. 2018. Multilingualism and European identification. *Sociological Inquiry* 88(3). 410–434.

Modiano, Marko. 1999. International English in the global village. *English Today* 15(2). 22–28.

Møller, Jannus & Norman Jørgensen. 2009. From language to languaging: Changing relations between humans and linguistic features. *Acta Linguistica Hafniensia* 41(1). 143–166.

Moran, William. 2002. *The amarna letters*. Baltimore: Johns Hopkins University Press.

Moreno, Frede & Lesley Ann Atilano-Tang. 2023. Hispanic colonial footprint and hubris in the Philippines: The case of Zamboanga city's creole Spanish-speaking population. Available at SSRN: https://ssrn.com/abstract=4545229 or https://doi.org/10.2139/ssrn.4545229.

Motschenbacher, Heiko. 2016. English in Europe and the postmodernist paradigm. In Andrew Linn (ed.), *Investigating English in Europe: Contexts and agendas*, 106–113. Berlin: De Gruyter.

Mugane, John. 2015. *The story of KiSwahili*. Athens, Ohio: Ohio University Press.

Mullen, Alex. 2011. Latin and other languages: Societal and individual bilingualism. In James Clackson (ed.), *A companion to the Latin language*, 527–548. Cambridge: Cambridge University Press.

O'Regan, John. 2021. *Global English and political economy*. Oxfordshire: Routledge.

Okombo, Patrick & Edwin Muna. 2017. The international status of Kiswahili: The parameters of Braj Kachru's model of world Englishes. *Africology: Journal of Pan African Studies* 10(7). 55–67.

Onysko, Alexander. 2016. Language contact and world Englishes. *World Englishes* 35(2). 191–195.

Ostler, Nicholas. 2006. *Empires of the word: A language history of the world*. New York: Harper Perrenial.

Ostler, Nicholas. 2011. *The last lingua franca. English until the return of Babel*. London: Penguin.

Patel, Mina, Mike Solly & Steve Copeland. 2023. *The future of English: Global perspectives*. London: British Council.

Pennycook, Alastair. 2007. *Global Englishes and transcultural flows*. Oxfordshire: Routledge.

Perez, M. Montero. 2022. Second or foreign language learning through watching audio-visual input and the role of on-screen text. *Language Teaching* 55(2). 163–192.

Pöll, Bernhard. 2017. Normes endogènes, variétés de prestige et pluralité normative. In Ursula Reutner (ed.), *Manuel des francophonies*, 65–86. Berlin: De Gruyter.

Pollock, Sheldon. 2009. *The language of the Gods in the world of men: Sanskrit, culture, and power in premodern India*. Berkeley: University of California Press.

Preisler, Bent. 1999. Functions and forms of English in a European EFL country. In Tony Bex & Richard Watts (eds.), *Standard English: The widening debate*, 239–267. Oxfordshire: Routledge.

Procházka, Stephan. 2021. Arabic dialectology. In Karin Ryding & David Wilmsen (eds.), *The Cambridge handbook of Arabic linguistics*, 214–243. Cambridge: Cambridge University Press.

Ramírez-Cruz, Héctor. 2020. Expressing uncontrolledness in narratives of accidental events: A comparison between native and non-native speakers of Spanish. *Lingua* 241. 102848.

Real Academia de la Lengua Española. 2010. *Ortografía de la Lengua Española*. Buenos Aires: Espasa.

Retsö, Jan. 2013. What is Arabic? In Jonathan Owens (ed.), *The Oxford handbook of Arabic linguistics*, 433–450. Oxford: Oxford University Press.

Reutner, Ursula. 2017. Vers une typologie pluridimensionnelle des francophonies. In Ursula Reutner (ed.), *Manuel des francophonies*, 9–64. Berlin: De Gruyter.

Riegelhaupt, Florencia, Roberto Luis Carrasco & Elizabeth Brandt. 2003. Spanish: A language of indigenous peoples of the Americas. In Jon Reyhner, Octaviana V. Trujillo, Roberto Luis Carrasco & Louise Lockard (eds.), *Nurturing native languages*, 129–140. Flagstaff, Arizona: Northern Arizona University.

Rjeoutski, Vladislav, Derek Offord & Gesine Argent. 2019. French as a diplomatic and official language in imperial Russia. *Jus Gentium: Journal of International Legal History* 4(2). 419–493.

Roberts, Peter. 2006. The odyssey of criollo. In Linda A. Thornburg & Janet M. Miller (eds.), *Studies in contact linguistics: Essays in honor of Glenn G. Gilbert*, 3–26. Lausanne: Peter Lang.

Rose, Heath & Nicola Galloway. 2019. *Global Englishes for language teaching*. Cambridge: Cambridge University Press.

Rose, Heath, Nicola Galloway & Jim McKinley. 2021. Global Englishes and language teaching: A review of pedagogical research. *Language Teaching* 54(2). 157–189.

Rovai, Francesco. 2021. On the standardisation of the Latin language and writing in the late republic. In Cesare Leta (ed.), *Studi classici e orientali LXVII*, 435–458. Pisa: Pisa University Press.

Ruppel, Antonia. 2017. *The Cambridge introduction to Sanskrit*. Cambridge: Cambridge University Press.

Saarinen, Taina. 2014. Language ideologies in Finnish higher education in the national and international context: A historical and contemporary outlook. In Anna Kristina Hultgren, Frans Gregersen & Jacob Thøgersen (eds.), *English in Nordic universities: Ideologies and practices*, 127–146. Amsterdam: John Benjamins.

Schneider, Edgar. 2003. The dynamics of new Englishes: From identity construction to dialect birth. *Language* 79(2). 233–281.

Seidlhofer, Barbara. 2001. Closing a conceptual gap: The case for a description of English as a lingua franca. *International Journal of Applied Linguistics* 11(2). 133–158.

Seidlhofer, Barbara. 2004. Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics* 24. 209–239.

Seidlhofer, Barbara. 2005. English as a lingua franca. *ELT Journal* 59(4). 339–341.

Seidlhofer, Barbara. 2021. English as a lingua franca in the European context. In A. Kirkpatrick (ed.), *The Routledge handbook of world Englishes*, 2nd edn, 389–407. London and New York: Routledge.

Siouffi, Gilles. 2010. *Le génie de la langue française. Études sur les structures imaginaires de la description linguistique à l'âge classique. Bibliothèque de Grammaire et de Linguistique 33*. Paris: Champion.

Smokotin, Vladimir, Anna Alekseyenko & Galina Petrova. 2014. The phenomenon of linguistic globalisation: English as the global lingua franca (EGLF). *Procedia - Social and Behavioral Sciences* 154. 509–513.

Song, Jieun & Paul Iverson. 2018. Listening effort during speech perception enhances auditory and lexical processing for non-native listeners and accents. *Cognition* 179. 163–170.

Spear, Thomas. 2000. Early KiSwahili history reconsidered. *International Journal of African Historical Studies* 33(2). 257–290.

Statista. 2004. Languages most frequently used for web content as of January 2024, by share of websites. https://www.statista.com/statistics/262946/most-common-languages-on-the-internet/ (accessed 04 05 2024).

Stewart, William. 1968. A sociolinguistic typology for describing national multilingualism. In Joshua Fishman (ed.), *Readings in the sociology of language*, 530–545. The Hague: Mouton.

Stotz, Peter. 2002. *Handbuch zur lateinischen Sprache des Mittelalters*, Vol. 1. Munich: C.H.Beck.

UNESCO. 1953. *The use of vernacular languages in education*. Paris: UNESCO.

van Putten, Marijn. 2020. Classical and modern standard Arabic. In Christopher Lucas & Stefano Manfredi (eds.), *Arabic and contact-induced change*, 57–62. Berlin: Language Science Press.

Versteegh, Kees. 2010. Pidgin Arabic and arabi sa'ab: The influence of the standard language in the history of Arabic. *Jerusalem Studies in Arabic and Islam* 37. 61–80.

Vila, Xavier. 2021. The hegemonic position of English in the academic field. Between scientific diglossia and academic lingua franca. *European Journal of Language Policy* 13(1). 47–73.

Wardhaugh, Ronald & Janet Fuller. 2021. *An introduction to sociolinguistics*. Hoboken: Wiley-Blackwell.

Watson, Janet. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Wenger, Etienne. 1998. *Communities of practice. Learning, meaning, and identity*. Cambridge: Cambridge University Press.

Wetzel, Mathis, Ekaterina Tskhovrebova, Pascal M. Gygax & Sandrine Zufferey. 2023. Pragmatic and syntactic constraints on French causal connectives: An evaluation of native and non-native speakers' sensitivity. *Journal of Pragmatics* 209. 89–107.

Widdowson, Henry. 1994. The ownership of English. *TESOL Quaterly* 28(2). 377–389.

Wilton, Antje. 2012. The monster and the zombie: English as a lingua franca and the Latin analogy. *Journal of English as a Lingua Franca* 1(2). 337–361.

Wright, Roger. 1982. *Late Latin and early romance in Spain and Carolingian France. (ARCA, classical and medieval texts, papers and monographs 8)*. Prenton: Francis Cairns Publications Ltd.

Wright, Roger. 1995. *Latin and the romance languages in the Early Middle Ages*. University Park: Pennsylvania State University Press.

Wright, Roger. 2003. *A sociophilological study of late Latin*. Turnhout: Brepols.

Wright, Roger. 2004. Latin and English as world languages. *English Today* 20(4). 3–13.

Wright, Sue. 2006. French as a lingua franca. *Annual Review of Applied Linguistics* 26. 35–60.

Wright, Roger. 2012. Convergence and divergence in world languages. In Juan Manuel Hernández-Campoy & Juan Camilo Silvestre (eds.), *The handbook of historical sociolinguistics*, 552–567. Hoboken: Blackwell Publishing.

Zamenhof, Ludwik. 1905. *Declaration of boulogne*. Boulogne-sur-Mer: Esperanto Congress.