



**HAL**  
open science

## VUM: a survey & corpus analysis of published journal articles covering the last ten years on the keywords ‘urban sociolinguistics’ & ‘multilingual communities’

Cédric Bruderemann, Jose Ignacio Aguilar Río, Myriam Abouzaid

### ► To cite this version:

Cédric Bruderemann, Jose Ignacio Aguilar Río, Myriam Abouzaid. VUM: a survey & corpus analysis of published journal articles covering the last ten years on the keywords ‘urban sociolinguistics’ & ‘multilingual communities’. 2020. hal-02543457

**HAL Id: hal-02543457**

**<https://hal.science/hal-02543457>**

Preprint submitted on 15 Apr 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **VUM: a survey & corpus analysis of published journal articles covering the last ten years on the keywords ‘urban sociolinguistics’ & ‘multilingual communities’**

Cédric BRUDERMANN<sup>1</sup>

Jose AGUILAR<sup>2</sup>

Myriam ABOUZAÏD<sup>2</sup>

<sup>1</sup> Sorbonne Université (College of Science and Engineering), Paris, France

<sup>2</sup> DILTEC, Sorbonne Paris Cité University, France

## **1. Introductory statement**

This paper aims to characterize the notion of *Vitality of Urban Multilingualism* (henceforth ‘VUM’) in the light of the latest developments in research and knowledge in applied linguistics. To do so, a corpus of scientific articles that span from 2006 to 2017 was built and analysed. In what follows, we first describe the collection protocol which was applied to (arbitrarily) create the corpus. We then propose a quantitative analysis of the data which was collected. A critical discussion about the way the corpus informs the concept of VUM is finally provided.

## **2. Introduction of the corpus: methodology and rationale behind our choices**

In order to put together a multilingual corpus on VUM which was expected to take into account several academic traditions, we first turned to three databases, i.e. ProQuest (English), Isidore (French) and Dialnet (Spanish). We agreed to concentrate on the three main working languages of these databases to conduct our documentary task, since they both made part of our common linguistic repertoire and allowed to scrutinize the notion of VUM from a variety of viewpoints and points of reference. Then, to query the databases, we used two phrases, which we translated in our three chosen target languages: English (“urban sociolinguistics” and “multilingual community”), French (“sociolinguistique urbaine” and “communauté plurilingue”) and Spanish (“sociolingüística urbana” and “comunidad plurilingüe”). We purposely chose these two phrases to explore the notion of VUM as it (arbitrarily) seemed to us that they could be considered as consistent – though not perfect – ways of paraphrasing it. The main weakness of such an approach is that by using (pseudo) synonyms – instead of the very notion itself – to query the databases, one can allegedly miss out on some specific nuances of the notion. In any case, this exploratory search led to hundreds of resulting matches. To narrow down the number of possibilities and, for more consistency, we decided to retain only the articles which were published in scientific journals (and which had incidentally gone through a double-blind peer review process) and for which the author(s)/editor(s) had supplied at least an abstract in English. In a limited number of cases (n=5), we translated into English the lists of keywords provided in French or Spanish by the authors themselves and we punctually (n=13) made up our own lists of English keywords, by selecting items in the provided abstracts in English. This way of proceeding led us to leave out 46 references and to gather a corpus of 125 references which we compiled within a Zotero collection<sup>1</sup>. We also created an online collaborative spreadsheet comprising 20 information categories that we defined together and further extracted from the abstracts the textual information that matched our requests. Out of the 20 categories that we initially created, seven sets of textometric data were eventually retained in this study (i.e. “year”, “publishing language”, “journal title”, “state/country”, “studied language(s)”, “keywords” and “abstract”), as the other ones were not yielding relevant information for our analyses. This data-mining procedure eventually allowed to conduct several statistical analyses.

### **3. Metadata search**

To conduct our survey, we first sought to generate metadata about our corpus. In this section, we present the metadata which was collected by using Libreoffice Calc (3.1 to 3.3) and Voyant Tools, “a web-based reading and analysis environment for digital texts”<sup>2</sup> (3.4 to 3.5).

#### **3.1. Years of publication**

1 Cf. <https://tinyurl.com/y7g4dhh3> (last visited on January, 29th, 2018)

2 Cf. <https://voyant-tools.org/> (last accessed January, 29th, 2018).

The corpus comprises studies which were published between 2006 and 2017. However arbitrary this choice of a decade may be, our aim has been to keep track of the latest traceable studies on VUM and, at the same time, to get closer to the period following the publication of the *Common European Framework of Reference for Languages* (henceforth “CEFR”, Council of Europe, 2001). We indeed regard the CERF as an important milestone and trace of the European linguistic policy, so much so that we believe its publication has had an impact on the way language policy is conceived of in Europe and elsewhere (Ishikawa, 2014), notwithstanding the adequacy of some critical standpoints to which we adhere unreservedly (Wisniewski, 2017). Figure 1 shows that the number of published papers dealing with VUM reaches a peak around the year 2014. A certain consistency seems to be observed in the number of published papers over the 2011-2015 period, when the mean value is 17,4 publications per year.

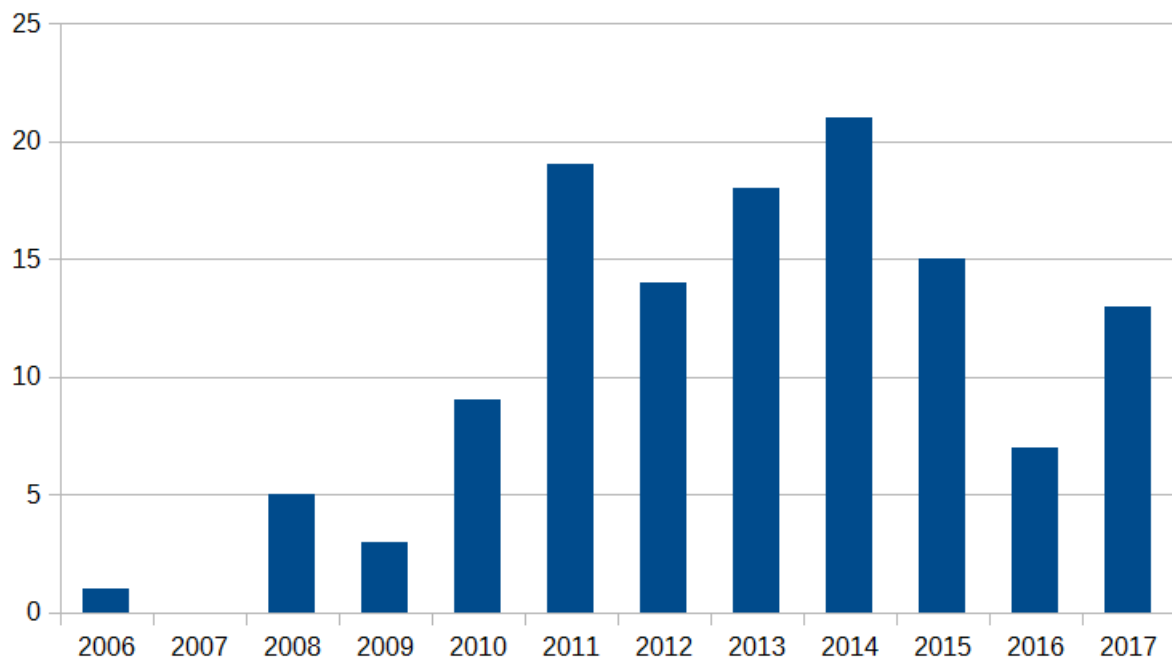


Figure 1. Distribution of the corpus items (by year of publication)

### 3.2. Publishing languages

The graph below shows the three languages in which the 125 references of our corpus were written.

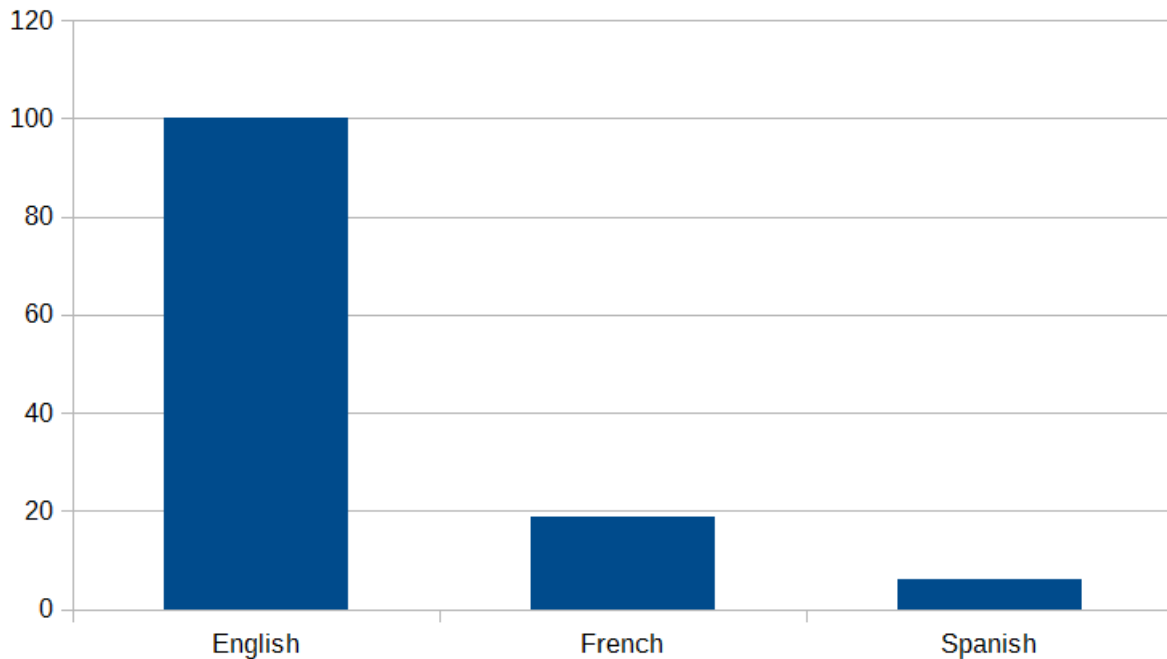


Figure 2. Publishing languages

English clearly stands out as the main publishing language (roughly 80 % of the corpus). This uneven distribution may be read as a consequence of English having become a default lingua franca in most scholarly arenas (Bitetti & Ferreras, 2016). However, it may also be the case that this observed distribution result from our scholarly reading habits, which may besides partly correlate to the aforementioned first consequence.

### 3.3. Journal titles

The graph below shows the diversity of journals (n = 50) from which the 125 references of our corpus were drawn, as well as their breakdown.

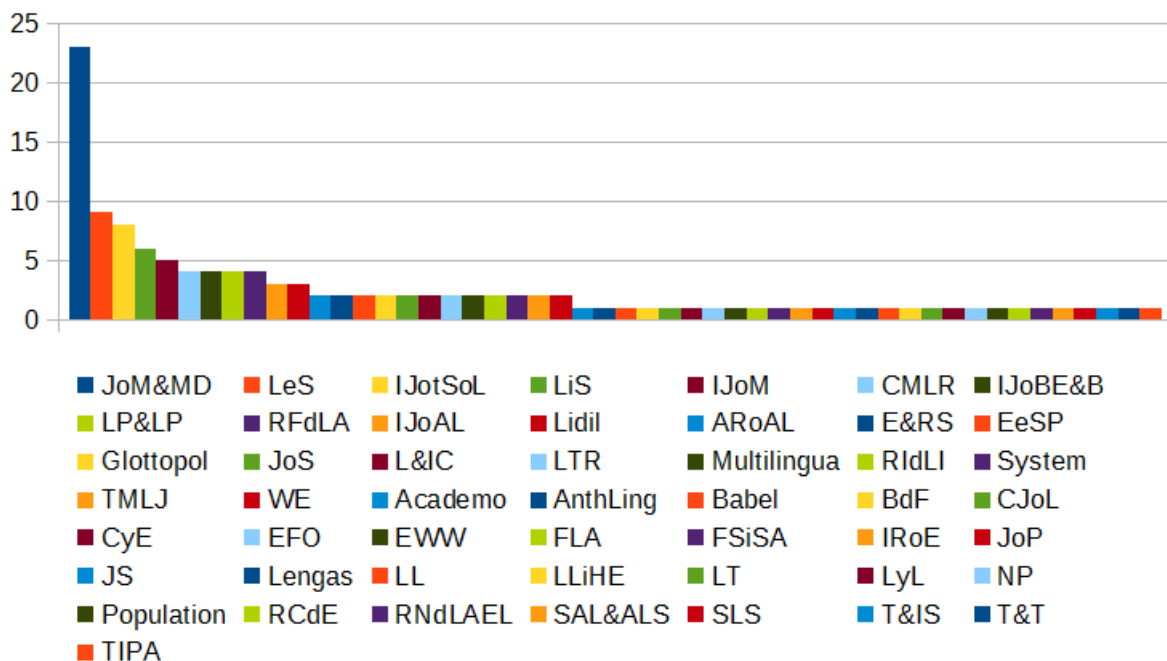


Figure 3. Distribution of the corpus items by journal title

The *Journal of Multilingual and Multicultural Development* (“JoM&MD”) contains the largest number of references in our corpus (n=23). 26 out of the 50 selected journals include only one reference, while 23 journals contain between two and four references. Finally, four journals – whose titles seem to be thematically related to the notion of VUM, such as the *Journal of Multilingual and Multicultural Development* – respectively count nine (*Langage et Société*, “LeS”), eight (*International Journal of the Sociology of Language*, “IJotSoL”), six (*Language in Society*, “LiS”) and five (*International Journal of Multilingualism*, “IJoM”) references.

### 3.4. States and countries

The map below shows the distribution of the geographical regions which were identified as contexts where fieldwork was conducted, as regards the 125 studies in our corpus.

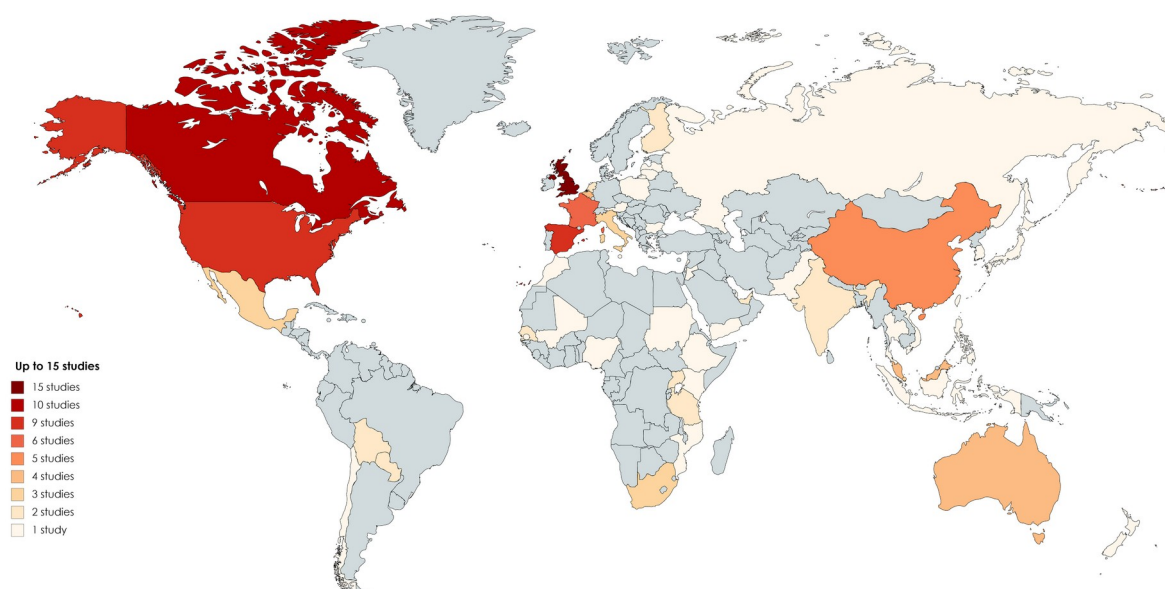


Figure 4. Distribution of the corpus items by geographical context

All the major geographical areas continents (but Antarctica) are represented in the corpus. Besides, three large geographical poles stand out in this survey, as they each account for, at least, five references in the corpus, namely: a) Central and North America; b) Western Europe; c) East Asia. As for the Middle East, South America and Africa, they appear to be more unevenly and scarcely represented.

### 3.5. Studied languages

The figure below shows the languages which were identified as research objects in our corpus and whose values are greater than or equal to two occurrences<sup>3</sup>.

<sup>3</sup> The full list of the studied languages is available at <http://tinyurl.com/j39634b> (last accessed January, 29th, 2018).

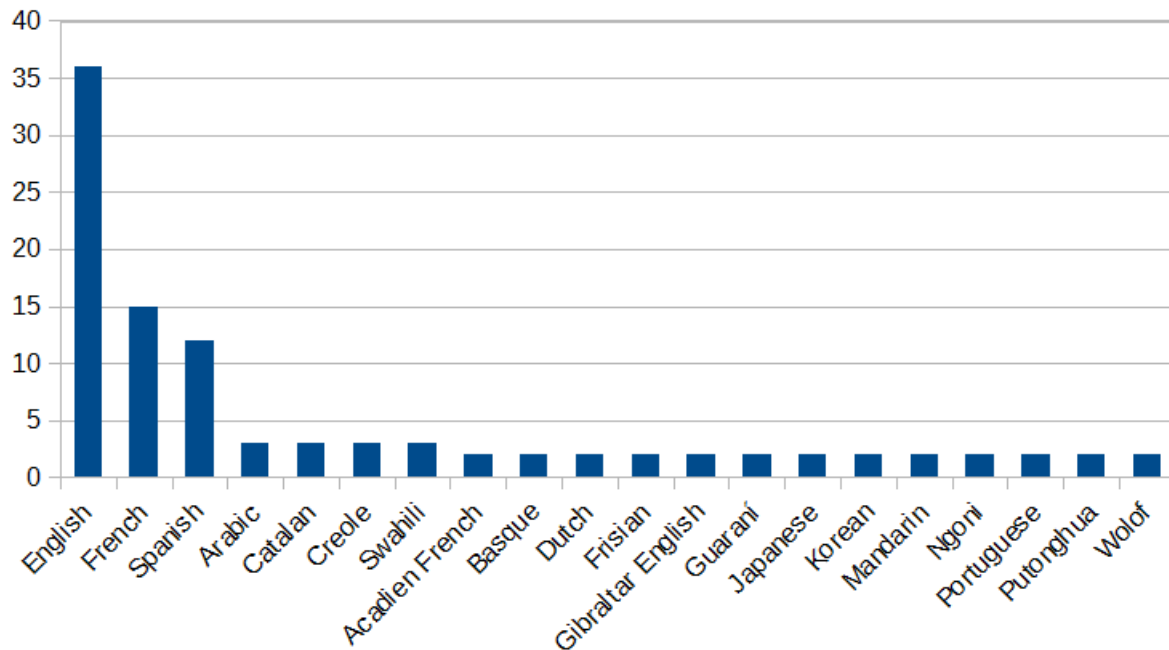


Figure 5. Studied languages (at least 2 occurrences)

English, French and Spanish feature the three languages that account for the greatest number of occurrences ( $n = 63$ ). There are more papers written in English ( $n = 100$ ) than articles that discuss the English language ( $n = 36$ ). Figure 5 above also shows the emergence of a first, populated, set of languages, which we will refer to as “minority languages” in what follows (such as “Guarani”, which is spoken in Paraguay, for instance). A second set of languages stands out, which corresponds to so-called “regional languages”, such as Basque, Catalan or Frisian. To refer to this set of languages, it can be noted that some authors use the expressions “local languages” (Conteh, 2012; Farr, 2011; Loakes *et al.*, 2013; Nyst, 2015; Tembe & Norton, 2008; Toffoli, 2015) or “small languages” (Gorter & Cenoz, 2011; Hancock, 2014; Loakes *et al.*, 2013; Toffoli, 2015) and that, outside the European boundaries, a greater diversity of expressions is used to refer to a similar sociolinguistic reality: “aboriginal language” and “heritage languages” to hint at the Kanakanavu and Saaroa languages in Taiwan (Liu, Chang, Li & Lin, 2015), “minority languages” for the Chinese context (Gao & Park, 2012), “vernacular languages”, to refer to the languages of the Solomon Islands (Jourdan & Angeli, 2014) and the expression “home languages”, to hint at the eleven official languages spoken in South Africa (Heugh, 2013).

The use of these various terms suggests the authors’ specific ideological standpoints – or their absence. Certain terms do not appear to possess a transnational scope. “Indigenous”, for instance, is extensively employed in the English and Spanish-speaking contexts, particularly in the Americas. Such is the term retained by Esteban-Guitart *et al.* (2015) concerning Chiapas (Mexico) or Campbell & Grondona (2010) for the research they conducted in Bolivia.

A further aspect in relation with these so-called “minority languages” concerns creoles and pidgins. Jourdan & Angeli (2014) account for a pidgin which is used as a lingua franca in the Solomon Islands, while other authors take an interest in the role(s) played by creole languages in educational contexts such as Chicago (Farr, 2011), Mauritius (Auleear Owodally & Unjore, 2013) or Saint Martin (Candau, 2015). Finally, as far as “minority languages” are concerned, some authors discuss research questions such as the ecology of languages, endangered languages and language vitality. Such is for instance the case in Jones’ (2008) account, who studies Jersey Norman French, or with Loakes *et al.* (2013), who focus on Walmajarri, an Austronesian language which is on the verge of dying out.

Exploring the ways in which the linguistic varieties are categorized leads to a third language set – which is reminiscent of Clyne’s (1992) “pluricentric” category – as the corpus indeed includes research questions that fit within a linguistic variation paradigm by virtue of which the functioning of local standards may be approached (cf. Bavoux *et al.*, 2008). Such studies about the emergence of new endogenous norms have been carried out by (for instance) Neumann-Holzschuh (2009), who studies Acadian and Louisiana French, Weston (2011), who focuses on Gibraltar English or Loakes *et al.* (2013), who present findings about Standard Australian English.

Eventually, among the 125 references, 36 deal with language learning and teaching in primary, secondary and higher educational contexts. These references revolve around four main, general, salient themes, which are presented in an alphabetical order in Table 1: a) first/heritage/second/X<sup>4</sup> language attitudes, development, maintenance and use (12 references); b) multilingual and multicultural policy implementation and attitudes to it (16 references); c) the parents’ attitudes, expectations and involvement vis-à-vis their children’s engagement in multicultural and multilingual schooling contexts (3 references); d) the teachers’ attitudes, beliefs and development regarding multilingual and multicultural policies and practice (5 references). These themes are neither exclusive, nor unique to each of the 36 references that have been identified as specifically dealing with educational matters. Rather, they appear to interweave and complete one another, yielding a complex characterisation of the diverse phenomena that multilingualism, understood as a human practice, may bring into play in different social situations, whether institutional or not. Table 1 below shows the contexts where these 36 educational-related studies were conducted, as well as the themes that we identified as chief for each of the 36 studies.

Study	Context	Identified theme
Adelin, 2013	Symmetric assessment of language skills in a nursery class in the multilingual context of La Réunion island.	Multilingual and multicultural policy implementation and attitudes to it
Afful, 2010	Ghanaian university students’ construction of their gendered identities through verbal behaviour	First/heritage/second/X language attitudes, development, maintenance and use
Armand, Lory &	Plurilingual creative expression	Multilingual and multicultural

4 We follow Dewaele (2017), who advocates the use of ‘X’, or ‘LX’, as a formal means to ideologically blur, or even do away with, whatever power relationships there may be among the languages within an individual’s repertoire, due to chronological, or any other kind of, organisation.



Rousseau, 2013	theatre workshops with underscholarised immigrant learning French in two welcoming classes in Montreal (Quebec)	policy implementation and attitudes to it
Aulear Owodally & Unjore, 2013	Mauritian Muslims' use and literacy ideologies concerning written Kreol	First/heritage/second/X language attitudes, development, maintenance and use
Burkhauser, Steele, Li, Slater, Bacon & Miller, 2016	Language proficiency assessment in the context of a dual-language immersion programme implemented in the public schools of the Portland area (Oregon, USA)	Multilingual and multicultural policy implementation and attitudes to it
Buttaro, 2014	The ineffectiveness of the English-only approach in educating Spanish-speaking students in the states of Arizona, California, Colorado, New Mexico and Texas (USA)	Multilingual and multicultural policy implementation and attitudes to it
Canagarajah & Ashraf, 2013	The implementation of multilingual educational policies in India and Pakistan	Multilingual and multicultural policy implementation and attitudes to it
Chevalier, 2017	The revival of a Sakha (Yakut) language education programme in the Republic of Sakha (Yakutia, Russian Federation).	Multilingual and multicultural policy implementation and attitudes to it
Conteh, 2012	Multilingual primary-aged children and their families in a post-industrial city in England	Parents' attitudes, expectations and involvement vis-à-vis their children's engagement in multicultural and multilingual schooling contexts
De Angelis, 2012	The effect of communication in the L1 or the L2 in the immediate living environment on the L1 and L2 acquisition process (both at a single point in time and overtime), in the multilingual context of South Tyrol (Italy)	First/heritage/second/X language attitudes, development, maintenance and use
Durán-Martínez et al., 2016	Novice and expert teachers' perceptions on four key dimensions of the CLIL approach in several primary and secondary schools in Spain	Teachers' attitudes, beliefs and development regarding multilingual and multicultural policies and practice
Early & Norton,	Fostering the development of	Multilingual and multicultural

2014	multilingual academic literacy in a rural Ugandan secondary school	policy implementation and attitudes to it
Gao & Park, 2012	Korean-Chinese parents' attitudes towards Putonghua and Korean in China	Parents' attitudes, expectations and involvement vis-à-vis their children's engagement in multicultural and multilingual schooling contexts
García & Sylvan, 2011	Newcomer immigrant students' plurilingual abilities development through heterogeneity, collaboration, learner-centeredness, language and content integration and responsibility at U.S. secondary schools for newcomer immigrants	Multilingual and multicultural policy implementation and attitudes to it
Garrett & Gallego Balsà, 2014	International and home students at the universities of Lleida (Catalonia, Spain) and Cardiff (Wales, UK)	First/heritage/second/X language attitudes, development, maintenance and use
Gorter & Cenoz, 2011	Language and education policy-making in the Spanish Basque Country and Friesland (The Netherlands)	Multilingual and multicultural policy implementation and attitudes to it
Gorter, 2015	Multilingual interactions in the Basque Country (Spain) and Friesland (The Netherlands)	Multilingual and multicultural policy implementation and attitudes to it
Gu & Benson, 2015	How teacher identities are discursively constructed and under the influence of a social structure in a teacher education course (Hong Kong and China)	Teachers' attitudes, beliefs and development regarding multilingual and multicultural policies and practice
Hancock, 2014	Language education policy as a lever for the promotion of Scotland's diverse languages	Multilingual and multicultural policy implementation and attitudes to it
Heugh, 2013	Multilingual education policy in South Africa	Multilingual and multicultural policy implementation and attitudes to it
Heugh, 2014	Inadequateness of multilingualism and multilingual education in Australia, as regards Indigenous language communities	Multilingual and multicultural policy implementation and attitudes to it
Izquierdo, García	Teachers' use of English and	Teachers' attitudes, beliefs and

Martínez, Garza Pulido & Aquino Zúñiga, 2016	Spanish in Mexican secondary-school classrooms	development regarding multilingual and multicultural policies and practice
Litzenberg, 2016	USA pre-service English language teachers' attitudes towards native/non-native Englishes in interaction	Teachers' attitudes, beliefs and development regarding multilingual and multicultural policies and practice
López-Gopar, Núñez-Méndez, Sughrua & Clemente, 2013	Translanguaging practices as the school norm, Oaxaca (Mexico)	First/heritage/second/X language attitudes, development, maintenance and use
Madariaga, Huguet & Janés, 2016	Immigrant students (with a great linguistic diversity)'s attitudes towards Catalan (Catalonia, Spain)	First/heritage/second/X language attitudes, development, maintenance and use
Marshall & Laghazaoui, 2012	French immersion graduates' language and literacy practices at an English-medium university in Vancouver, Canada	Multilingual and multicultural policy implementation and attitudes to it
Milans, 2010	Social and institutional construction of culture at school (Madrid (Spain) and Zhejiang (China))	First/heritage/second/X language attitudes, development, maintenance and use
Ngcobo, 2014	Language identity and bi/multilingual education in South Africa	Multilingual and multicultural policy implementation and attitudes to it
Nguyen & Hamid, 2016	A qualitative study of Vietnamese ethnic minority students' language attitudes, identity and L1 maintenance	First/heritage/second/X language attitudes, development, maintenance and use
Oriyama, 2012	Japanese literacy among school-age children of Japanese heritage in Sydney, Australia	First/heritage/second/X language attitudes, development, maintenance and use
Prasad, 2012	Language and multicultural policies in Canada as political frames that limit language rights and schooling for allophone immigrant children	Multilingual and multicultural policy implementation and attitudes to it
Pulinx & Avermaet, 2014	Interaction between language ideologies, education policies and teacher beliefs about monolingualism in Flanders (Belgium)	Teachers' attitudes, beliefs and development regarding multilingual and multicultural policies and practice

Riches & Curdt-Christiansen, 2010	Exploration of the parents' aspirations, expectations, support and involvement in their children's education in two different linguistic and ethnic communities in Montréal, Canada	Parents' attitudes, expectations and involvement vis-à-vis their children's engagement in multicultural and multilingual schooling contexts
Tembe & Norton, 2008	Community responses to the language education policy in primary schools in rural and urban communities in eastern Uganda	First/heritage/second/X language attitudes, development, maintenance and use
Toffoli, 2015	Students' plurilingual profiles, attitudes and language learning preferences, University of Strasbourg	First/heritage/second/X language attitudes, development, maintenance and use
Wei, 2011	Multilingual and multimodal practices of British Chinese children attending various complementary schools in 3 English cities	First/heritage/second/X language attitudes, development, maintenance and use

---

Table 1. 36 education-related studies within our corpus

## 4. Qualitative analyses

### 4.1. Keywords

In order to identify what the most frequent entries in the corpus were, the keywords of all the selected abstracts were brought together and classified in alphabetical order to form a corpus of keywords (“CK”). In order to also take the collocations into account in this analysis (“language contact”, for example), the spaces between the words were deleted<sup>5</sup>. We eventually used a piece of software called KWIC to carry out a lexicometric analysis of the CK. The analysis revealed that the CK comprised 736 tokens and 561 types<sup>6</sup>. In this survey, while 449 items were to be found once, some of them were repeated several times. As such, the ten most frequent entries of the CK are presented in figure 6 below<sup>7</sup>.

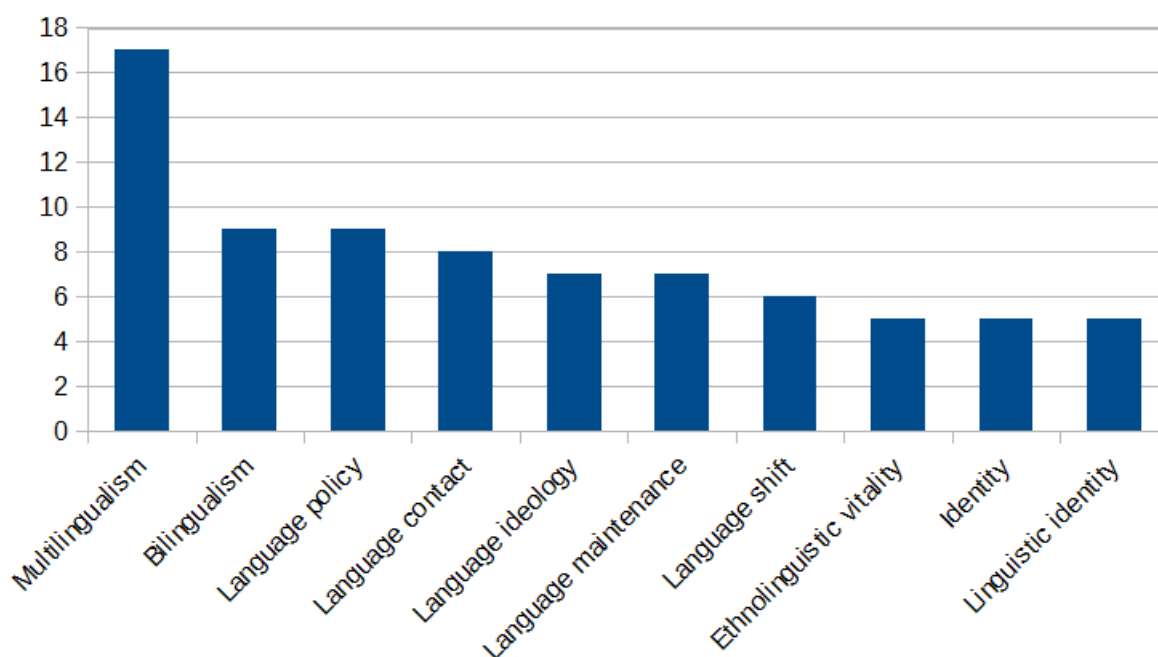


Figure 6. Top ten keywords in the CK

### 4.2. Depicting trends

Our objective is to characterize the concept of VUM. In this respect, the most frequent keywords seem to reveal that what shall be understood by VUM is the “ethnolinguistic vitality” (item #8, figure 6 above) of the languages practices underway in today’s multilingual and multicultural urban settings.

5 In this analysis, we have kept the distinction between the singular and plural forms of the entries. As such, pairs like “*Languageattitude/Languageattitudes*” or “*Culturalidentities/Culturalidentity*” were for instance considered as two different entries.

6 ‘Token’ refers to the total number of words in a text or a corpus, regardless of how often they are repeated. As for ‘type’, it refers to the number of distinct words to be found in a text or a corpus.

7 The full list of keywords is available, at a spreadsheet file, at <https://tinyurl.com/jy6t9cb> (last accessed January, 29th, 2018).

On the other hand, the corpus seems to highlight that ‘urban sociolinguistics’ comes with specific features:

1. Depending on the urban contexts, a quantitative distinction has to be drawn between ‘multilingualism’ and ‘bilingualism’ (items #1 and #2, figure 6 above). The “languages in contact” (item #4) in a given urban place – which denote linguistic repertoires possibly made up of two or more languages – are a context-specific variable which, in turn, (i) contributes to the shaping of one’s “linguistic identity” (item #10, figure 6 above) and (ii) influences the city dwellers’ language practices (i.e. “language maintenance and shift”, items #6 and #7, figure 6 above).
2. Besides these cultural and identity-related challenges (item #9, figure 6 above), the concept of VUM also seems to entail a strong political dimension, i.e. the efforts (“language policy”, item #3, figure 6 above) brought forward by governments to regulate the linguistic practices underway within the limits of their territories. As such, in order to perpetuate their language policies, governments seem to resort to language ideologies (item #7, figure 6 above) and to use specific institutional places to implement their policies.
3. One lever which seems to be relied upon by governments to achieve this objective is education. Although not ranking amongst the most common entries in this survey, the semantic field that revolves around the educational sector is indeed very present in the corpus (cf. table 1 above). And even though this presence is ‘plural’ – as it is embodied by the use of numerous synonyms – it nonetheless remains consistent throughout in the CK, with phrases such as ‘bilingual education’, ‘multilingual education’, ‘language education policy’, ‘mother tongue education’, ‘trilingual education’, ‘language education policy’, ‘multilingual educational policies’ or ‘teacher education’. This plurality therefore seems to suggest a strong link between VUM and (language) education ; it is also the sign that the sociolinguistic contexts are numerous, that there is not one way to implement a language education programme and that the language dynamics/changes which are underway in today’s urban settings require to establish (language) educational policies on a case-by-case basis, in order to respect and adapt to the cultural and linguistic features of the contexts, whether they be monolingual, bilingual or multilingual and whether they include actual speakers or signers.

#### 4.3. Corpus of abstracts

In a second phase, all the abstracts were compiled in another single document, in order to further build a corpus of abstracts (CA)<sup>8</sup>. The CA comprises 21688 tokens and 3,935 types (among which 2,059 appear once). When we performed our statistical analysis, the spaces between the words were kept. After removing the stopwords, the items whose number of apparitions was greater than or equal to 20 were the following:

<sup>8</sup> Available at <http://tinyurl.com/j39634b> (last accessed January, 29th, 2018). The corresponding Voyant-Tools project is available at <https://tinyurl.com/y9wl4wxl> (last accessed January 29th, 2018). Please note that Voyant-Tools and KWIC process the data in slightly different ways. Some of the analyses built-in the Voyant-Tools interface may thus differ from those we have presented in this article and which we obtained via KWIC. Notwithstanding these slight divergences, we believe our readership may find the Voyant-Tools project interesting and worth reading.

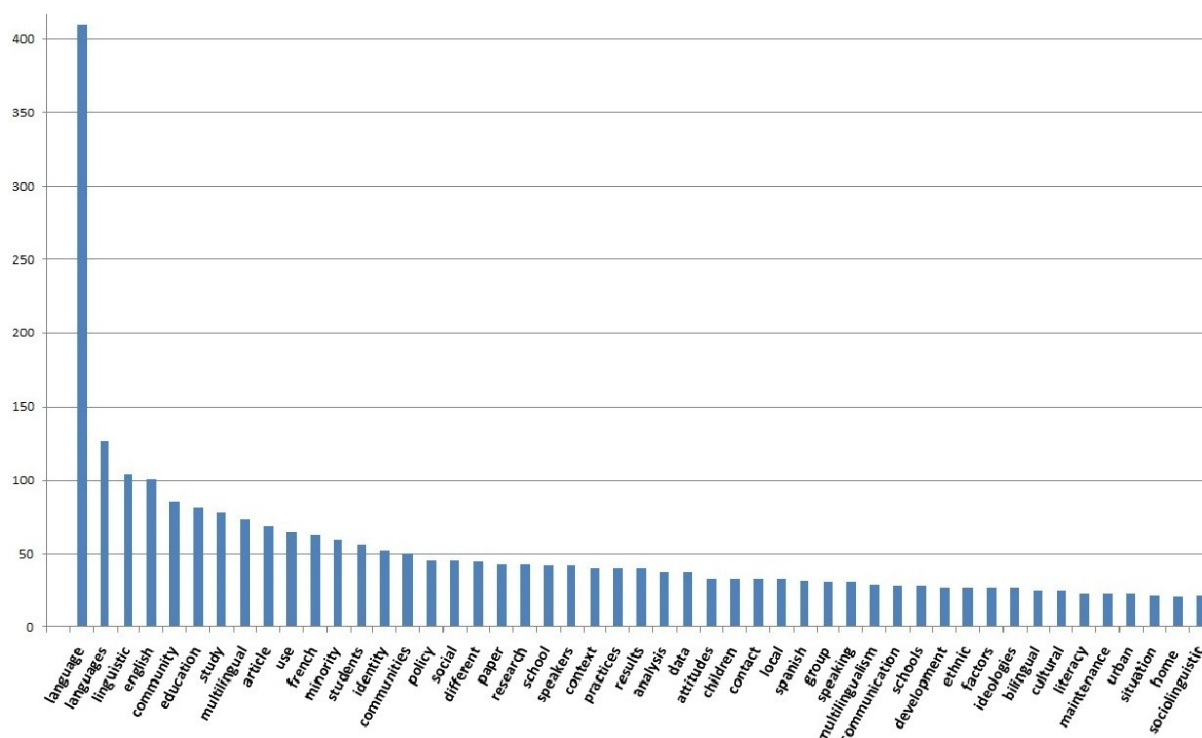


Figure 7. Most frequent terms in the CA

#### 4.4. Interpreting the data

For this analysis, we first classified the CA's most frequent entries (n=50) in different categories, among which two proved irrelevant to inform the concept of VUM:

1. The entries which are related to the academic sphere and academic writing. This is coherent with the kind of "literature" the abstracts were drawn from: study (78); article (69), paper (43); research (43); analysis (37); data (37); results (40).
2. The entries related to the main topics tackled in the selected abstracts: language (409); languages (126); linguistic (104).

The other categories highlighted that the distinction between one's private (identity (52); group (31); ethnic (27)) and public (policy (46); education (82); ideologies (27)) spheres could also be noted in this corpus, with some new details allowing to refine our first analysis (cf. 4.2.):

1. The notion of 'coexistence' – i.e community/ies (86/51) with different context (40)-specific (local (33)) language practices (40), living in areas where a given language is predominant (in particular, English (101), French (63) and Spanish (32)) and is in contact (33) with other languages – can now quantitatively be informed: urban (23) multilingualism (29) can refer both to bilingual (25) and multilingual (73) settings but, if we refer to the number of entries, urban settings tends to be more characterized by multilingual than bilingual (multilingual (73) + multilingualism (29) = 102 vs. bilingual (25)) language practices.

2. The notion of ‘contact’, which refers both to the communication (28) process itself and to the language choices (use (65)) made by multilingual speakers (42) when they interact with ‘different’ (45) people, that is people with whom they do not exactly share the same linguistic repertoires. The linguistic adjustments which become necessary in multilingual settings to foster mutual understanding and communication thus affect the speakers’ communicative attitudes (33) and language practices (40). Besides, the social context where the language interactions are taking place (at school (42 + 28 for ‘schools’), within the community (86 + 51 for ‘communities’), at home (21)) seem to be at the core of an ‘adaptive’ process (i.e. language maintenance (23) or switch, for instance) which could be qualified as ‘language dynamics’.
3. The relationship between VUM and language education. To explore the meaning underlying this association, the most frequent terms in the CA were relied upon. If these (unsurprisingly) reveal that schools (28) are settings which are characterized by regular encounters between teachers (20) and children (33), they also highlight that, in this very particular context, children become students (56), i.e. “individual-learners”, who correspond to complex and multifaceted entities composed of social, emotional, cognitive and discourse-based variables and whose singularities embrace features as varied as their own particular identities, personalities or cognitive styles (Aguilar & Brudermann, 2014). In turn, schools make up social (46) situations (22) of a very particular kind as (i) they call into play both the learners’ public (formal lessons, activities, lectures, etc) and private spheres (socialization with friends, informal activities, recreational activities, breaks, etc), whether it be in intertwined (a school trip, for example) or alternate ways ; (ii) educational institutions are also led to organize activities which imply the participation of individuals from diverse cultural (25) backgrounds and who need to collaborate in at least one official language of instruction (and, most of the time, over extended periods of time) to achieve the objectives which are assigned to them. This very particular combination of factors – in which speakers from minority (59) and majority groups form one whole – thus contributes to the construction of a common reality (27) and social cohesion. Educational policies therefore include a strong sociolinguistic (22) dimension (ie the use of at least one common means of communication), whether it be for speaking (31) or writing, through literacy (23) development (27).

A parallel can thus be drawn between the analyses performed on the data extracted from the CK and the CA. In order to further refine them, a collocation analysis of the CA was performed.

#### 4.5. Making collocations meaningful

In this study, it was not realistic to rely on un-spaced words to perform a statistical analysis, given the size of the corpus. However, in order to further explore the concept of VUM, we sought to conduct a collocation search to identify the words with which the most frequent entries<sup>9</sup> of the CA (Figure 7) were most commonly associated, using KWIC.

<sup>9</sup>All but the items that were related to the academic sphere and academic writing.



For each analysis, the results had to be sorted out, as some expressions were meaningless, did not exist or were not recurrent enough in the corpus and, therefore, not significant<sup>10</sup>. All the significant collocations were then compiled in one single document<sup>11</sup>. At the end of the analysis, the number of collocations per entry varied between 64 (for the entry ‘language(s)’) and 1 (‘factors’, ‘children’, ‘ideologies’, etc.) expressions. No collocations were besides highlighted for the entries ‘*situation*’, ‘*maintenance*’, ‘*literacy*’, ‘*sociolinguistic*’ and ‘*urban*’ of the CA.

#### 4.6. Interpreting the data

In line with what was previously outlined, this collocation extraction seems to indicate that VUM can be understood as a twofold concept:

1. One that refers to the ‘free’ language practices which are underway within specific cultural communities or in one’s private sphere (‘home language’) and which can be characterized by the use of ‘minority languages’ (compared to the ‘dominant’ and ‘official language(s)'). This concept of ‘minority languages’ is, in turn, twofold: it can either refer to ‘local’, ‘regional’ or ‘heritage’ languages, when the speakers are also native-born speakers of a country and part of a community which has long been established in this country, or ‘foreign’, ‘second’ or ‘migrant’ languages, when the speakers are part of a ‘migrant community’ (i.e. diaspora, asylum seekers, refugees, foreign workers, etc.) which form ‘minority’ (and sometimes ‘second-generation’) ‘ethnic’ groups. In both cases, the spoken languages are ‘different’ from the ‘official’, ‘national’ and dominant languages which (among other things) embody the public sphere. As far as VUM is concerned, the entries ‘use’, ‘groups’ and ‘linguistic’ reveal that ‘linguistic vitality’ in multilingual contexts accounts for the ‘linguistic diversity’ of the different ‘linguistic communities’ living within one shared territory. As these ‘language groups’ seem to have ‘local’ or ‘community’ uses of their languages – i.e. language practices which do not extend beyond the limits of their language groups – these ‘language uses’ seem ‘limited’, when we look at language use from a broader perspective, that of the nation as a whole (i.e. the public sphere and the ‘public use’ and ‘national use’ of one or several language(s));

10 The arbitrary threshold of ‘significance’ we set in this study was greater than or equal to five occurrences of a given collocation.

11 Available at <https://tinyurl.com/ycde53xo> (last accessed on January, 29th, 2018). In this document, the matching entries either precede or come after the keywords (highlighted in yellow).

2. One that is related to the language contact(s) between members of different communities within shared urban areas and, to a greater extent, shared countries. This part of the definition is therefore closely related to the regulatory efforts brought forward by governments to (i) foster interactions between communities (more particularly in official places such as courts, schools, the administration, etc.) and (ii) to further allow the nation as a whole to be able to live together and to build a shared future together. To meet such objectives, as it was illustrated above, drafting ‘language education’ schemes seems to be a necessity in ‘multilingual contexts’ and, depending on the urban contexts, these policies may aim at fostering either a ‘bilingual’ or a ‘multilingual education’, particularly in strategic educational places such as ‘secondary schools’, ‘universities’, ‘language schools’, ‘primary schools’, ‘community schools’ or ‘public school’. In turn, the ‘multilingual development’ that emerges from such policies is bound to favour the emergence of ‘multilingual communities’ characterized by ‘multiple/plural identities’ that encompass ‘community’, ‘local’ and ‘urban’ dimensions. As such, in the public sphere, cultural diversity and the encounter with the Other (i.e. different ‘ethnic groups’) seems to be the driving force behind ‘contemporary language multilingualism’ as it offers fertile ground for cultural hybridization and intercultural communication.

## 5. Discussion

In this study, we have decided to adopt an inclusive approach and give the users of the languages dealt with in the corpus their legitimate place, as far as multilingualism as a social practice and a human reality is concerned. As a consequence, a comprehensive characterisation of multilingual vitality in urban areas throughout the world has been attempted, to which we have systematically referred as ‘VUM’. It is our hope that our analyses may come across as an accurate blueprint of “the VUM complexity”.

There are nonetheless evident limits to this survey, which mainly originate from the choices we have made. The first limit has to do with the limited scope that one decade represents as a potential and empirical time measure wherefrom to spot socially, human-related research objects. Socially, empirical readings of human beings take time, most often years, or even decades (Tang, 2010; Raghavan, 2014). It is hard for the social sciences to be in phase – let alone to keep up with – the pace of the naturally occurring, social and, human-based phenomena<sup>12</sup>. Major human, linguistic and social events have taken place since 2006, such as the independence of Kosovo and South Sudan (in 2008 and 2011 respectively), the annexation of Crimea in 2014, Brexit, the dismantling of the Calais refugee settlement and the fall of Aleppo, in 2016<sup>13</sup>. Research underway and research to come in applied linguistics will delve into the causes of these major events, and certainly suggest likely consequences as far as language related phenomena are concerned.

<sup>12</sup> The “13 novembre” research project, which follows the series of attacks made on Greater Paris in November 2015, appears to us as a remarkable exception (cf. <http://www.memoire13novembre.fr/>, last accessed January, 29th, 2018).

<sup>13</sup> Among the 125 research articles that make up our corpus, not one is concerned about either of the aforementioned events.

A second limit revolves around the categories in our corpus whose analyses we have not pursued. Since our objective has been to set the necessary tools and data to conduct a content analysis of sorts, we have focused on seven categories that we perceived as relevant to inform the concept of VUM (cf. §3.1 through §4.6). Among those categories we have not analysed, ‘theoretical framework’ and ‘methods’ stand out. Of course, as empirical disciplines, research in the humanities and social sciences requires to rely upon methods and theories. Similar phenomena may lead to different readings according to the standpoint chosen by the research teams. All science that is conducted by humans (and the social sciences are no exception to this) is conducted by individuals who carry attitudes, beliefs, biases and ideologies. Consequently, as social scientists aim at fathoming aspects of human phenomena, they may inadvertently induce others to adhere to specific, maybe unconscious, attitudes towards these very phenomena. Of course, the present authors are no exception to this either.

In any case, the present survey on VUM may be understood as a somewhat new skin – at least not a very old one – to convey a research construct that is certainly not new as far as applied linguistics and sociolinguistics are concerned (Calvet, 1999)<sup>14</sup>. VUM points at human practices that may very well have been existing ever since communities using different codes to make sense of reality are aware of the existence of each other. And getting to grasp elements about the intricate interplay between language(s) and attitude(s) across the globe may lead some of us to better apprehend the complexity of our human societies or, at the very least, gain some perspective as to some aspects of what appears to be common to the latter.

## **6. Conclusion**

<sup>14</sup> And neither has it been our intention to dispute this.

Whatever the social contexts tackled in our selected abstracts, our analyses seem to reveal that the major political trend in multicultural urban settings is to foster social inclusion – so as to further promote peace and stability – and that, to achieve such a goal, education and the use of at least one official language are tools of choice. Whatever the educational place, social inclusion can indeed be addressed in the curricula both theoretically and practically, by promoting workshops/lectures/activities/trips that for instance entail mutual respect, collaboration, interacting or opening up to the world ; and these can only be achieved through the use of one common tool of cohesion/means of communication among participants who may as well have completely different culturo-linguistic backgrounds. As such, it is in the interest of the states to strive to ensure appropriate conditions for the coexistence of people and of human life in the public sphere and for the individuals (i) not to be discriminated against for being part of a minority group (ii) and to be able to pursue the language practices they wish in their private spheres. Such outcomes are common in countries subscribing to democratic values. However, if group cohesiveness (as a whole) – which highly depends on the ‘quality’ of the relationships between the various social communities in contact – is of utmost importance for developing inclusive social environments (whether urban, suburban or rural), History shows that other political projects, pursuing other objectives can also be implemented in multicultural settings and that such projects may, in turn, affect languages<sup>15</sup>. The social construct ‘multilingual’ is therefore subjected to forms of ideology which may be connected with lexical choices and which also account for language vitality, particularly as, as our analyses seem to show it, this expression can be considered as a suitable synonym for ‘language dynamics’, i.e. the linguistic strategies behind language practice (e.g. language maintenance/shift, code switching, etc.) in multilingual social contexts.

## 7. References

- Aguilar Río, J. I., & Brudermann, C. (2014). Language learner. In C. Fäcke (Ed.), *Manual of Language Acquisition* (pp. 291–307). Berlin, Boston: De Gruyter Mouton. Retrieved from <http://www.degruyter.com/view/product/186395>
- Afful, J. B. A. (2010). Address forms among university students in Ghana: a case of gendered identities? *Journal of Multilingual and Multicultural Development*, 31(5), 443–456. <https://doi.org/10.1080/01434632.2010.505655>
- Armand, F., Lory, M.-P., & Rousseau, C. (2013). Ateliers d’expression théâtrale plurilingues en classe d’accueil. *Lidil*, (48), 37–55.
- Aulear Owodally, A. M., & Unjore, S. (2013). Kreol at school: a case study of Mauritian Muslims’ language and literacy ideologies. *Journal of Multilingual and Multicultural Development*, 34(3), 213–230. <https://doi.org/10.1080/01434632.2013.763813>
- Bavoux, C., & Prudent, L.-F. (2008). *Normes endogènes et plurilinguisme: aires francophones, aires créoles*. ENS Editions.
- Bitetti, M. S. D., & Ferreras, J. A. (2016). Publish (in English) or perish: The effect on citation rate of using languages other than English in scientific publications. *Ambio*, 1–7. <https://doi.org/10.1007/s13280-016-0820-7>
- Burkhauser, S., Steele, J. L., Li, J., Slater, R. O., Bacon, M., & Miller, T. (2016). Partner-Language Learning Trajectories in Dual-Language Immersion: Evidence From an Urban District. *Foreign Language Annals*, 49(3), 415–433. <https://doi.org/10.1111/flan.12218>

<sup>15</sup> For instance, following the war, the communities of what was once known as Yugoslavia wished to live apart from one another and the once single common language formerly known as ‘Yugoslavian’ has now been turned into a “mosaic of languages” – i.e. Serbian, Croatian, Bosnian, and Montenegrin (a linguistic process called ‘ausbau languages’) – for the members of the communities who are in touch in this part of the world to both better outline their own specific cultural features and distinguish themselves from the others.

- Buttaro, L. (2014). Effective Bilingual Education Models. *Revue Française de Linguistique Appliquée*, XIX(2), 29–40.
- Calvet, L.-J. (1999). Aux origines de la sociolinguistique la conférence de sociolinguistique de l'UCLA (1964). *Langage et société*, 88(1), 25–57. <https://doi.org/10.3406/lsoc.1999.2867>
- Campbell, L., & Grondona, V. (2010). Who speaks what to whom? Multilingualism and language choice in Misiones La Paz. *Language in Society*, 39(5), 617–646. <https://doi.org/10.1017/S0047404510000631>
- Candau, O.-S. (2015). Pour une étude du parler à Saint-Martin : l'exemple des pratiques verbales de lycéens plurilingues. *TIPA. Travaux interdisciplinaires sur la parole et le langage*, (31). <https://doi.org/10.4000/tipa.1501>
- Canagarajah, S., & Ashraf, H. (2013). Multilingualism and Education in South Asia: Resolving Policy/Practice Dilemmas. *Annual Review of Applied Linguistics*, 33, 258–285. <https://doi.org/10.1017/S0267190513000068>
- Chevalier, J. F. (2017). School-based linguistic and cultural revitalization as a local practice: Sakha language education in the city of Yakutsk, Russian Federation. *Nationalities Papers*, 45(4), 613–631. <https://doi.org/10.1080/00905992.2016.1275534>
- Clyne, M. (1991). *Pluricentric Languages: Differing Norms in Different Nations*. Walter de Gruyter.
- Conseil de l'Europe. (2001). *Un Cadre Européen de Référence pour les Langues: Apprendre, Enseigner, Évaluer*. Strasbourg: Division des Politiques Linguistiques.
- Conteh, J. (2012). Families, pupils and teachers learning together in a multilingual British city. *Journal of Multilingual and Multicultural Development*, 33(1), 101–116. <https://doi.org/10.1080/01434632.2011.638077>
- De Angelis, G. (2012). The Effect of Population Distribution on L1 and L2 Acquisition: Evidence from the Multilingual Region of South Tyrol. *International Journal of Multilingualism*, 9(4), 407–422. <https://doi.org/10.1080/14790718.2012.714383>
- Dewaele, J.-M. (2017). Why the Dichotomy 'L1 Versus LX User' is Better than 'Native Versus Non-native Speaker'. *Applied Linguistics*, amw055. <https://doi.org/10.1093/applin/amw055>
- Durán-Martínez, R., Beltrán-Llavador, F., & Martínez-Abad, F. (2016). A contrastive analysis between novice and expert teachers' perceptions of school bilingual programmes. *Cultura y Educación*, 28(4), 738–770. <https://doi.org/10.1080/11356405.2016.1237339>
- Early, M., & Norton, B. (2014). Revisiting English as medium of instruction in rural African classrooms. *Journal of Multilingual and Multicultural Development*, 35(7), 674–691. <https://doi.org/10.1080/01434632.2014.908890>
- Esteban-Guitart, M., Viladot, M. À., & Giles, H. (2015). Perceived institutional support among young indigenous and Mestizo students from Chiapas (México): a group vitality approach. *Journal of Multilingual and Multicultural Development*, 36(2), 124–135. <https://doi.org/10.1080/01434632.2014.898645>
- Farr, M. (2011). Urban plurilingualism: Language practices, policies, and ideologies in Chicago. *Journal of Pragmatics*, 43(5), 1161–1172. <https://doi.org/10.1016/j.pragma.2010.10.008>
- García, O., & Sylvan, C. E. (2011). Pedagogies and Practices in Multilingual Classrooms: Singularities in Pluralities. *The Modern Language Journal*, 95(3), 385–400. <https://doi.org/10.1111/j.1540-4781.2011.01208.x>
- Gao, F., & Park, J. (2012). Korean-Chinese parents' language attitudes and additive bilingual education in China. *Journal of Multilingual and Multicultural Development*, 33(6), 539–552. <https://doi.org/10.1080/01434632.2012.692683>

- Gorter, D. (2015). Multilingual interaction and minority languages: Proficiency and language practices in education and society. *Language Teaching*, 48(1), 82–98. <https://doi.org/10.1017/S0261444812000481>
- Gorter, D., & Cenoz, J. (2011). Multilingual education for European minority languages: the Basque Country and Friesland. *International Review of Education*, 57(5), 651–666. <https://doi.org/10.1007/s11159-011-9248-2>
- Gu, M. (Michelle), & Benson, P. (2015). The formation of English teacher identities: A cross-cultural investigation. *Language Teaching Research*, 19(2), 187–206. <https://doi.org/10.1177/1362168814541725>
- Hancock, A. (2014). Language education policy in multilingual Scotland: opportunities, imbalances and debates. *Language Problems and Language Planning*, 38(2), 167–191. <https://doi.org/10.1075/lplp.38.2.04han>
- Heugh, K. (2013). Multilingual Education Policy in South Africa Constrained by Theoretical and Historical Disconnections. *Annual Review of Applied Linguistics*, 33, 215–237. <https://doi.org/10.1017/S0267190513000135>
- Heugh, K. (2014). Turbulence and dilemma: implications of diversity and multilingualism in Australian education. *International Journal of Multilingualism*, 11(3), 347–363. <https://doi.org/10.1080/14790718.2014.921180>
- Ishikawa, F. (2014). Enjeux de la contextualisation du CECRL dans un pays extra-européen : le cas de l'enseignement du FLE au Japon. In S. Y.-R. & E. S. M. Derivry-Plard G. Alao (Ed.), *Dispositifs éducatifs en contexte mondialisé et didactique plurilingue et pluriculturelle* (pp. 149–164). Peter Lang. <https://doi.org/10.3726/978-3-0352-0266-3>
- Izquierdo, J., García Martínez, V., Garza Pulido, M. G., & Aquino Zúñiga, S. P. (2016). First and target language use in public language education for young learners: Longitudinal evidence from Mexican secondary-school classrooms. *System*, 61, 20–30. <https://doi.org/10.1016/j.system.2016.07.006>
- Jones, M. C. (2008). Identity Planning in an Obsolescent Variety: The Case of Jersey Norman French. *Anthropological Linguistics*, 50(3/4), 249–265.
- Jourdan, C., & Angeli, J. (2014). Pijin and shifting language ideologies in urban Solomon Islands. *Language in Society*, 43(3), 265–285. <https://doi.org/10.1017/S0047404514000190>
- Liu, D. T., Chang, Y., Li, P. J., & Lin, J. (2015). Language shift of Taiwan's indigenous peoples: a case study of Kananavu and Saaroa. *Journal of Multilingual and Multicultural Development*, 36(7), 729–749. <https://doi.org/10.1080/01434632.2015.1022179>
- Litzenberg, J. (2016). Pre-service teacher perspectives towards pedagogical uses of non-native and native speech samples: Pedagogical uses of non-native and native speech samples. *International Journal of Applied Linguistics*, 26(2), 168–189. <https://doi.org/10.1111/ijal.12084>
- Loakes, D., Moses, K., Wigglesworth, G., Simpson, J., & Billington, R. (2013). Children's language input: A study of a remote multilingual Indigenous Australian community. *Multilingua*, 32(5), 683–711. <https://doi.org/10.1515/multi-2013-0032>
- Neumann-Holzschuh, I. (2009). Contact-induced structural change in Acadian and Louisiana French : mechanisms and motivations. *Langage et Société*, (129), 47–68.
- Ngcobo, S. (2014). The Struggle to Maintain Identity in Higher Education among Zulu-Speaking Students. *International Journal of Bilingual Education and Bilingualism*, 17(6), 695–713. <https://doi.org/10.1080/13670050.2014.953776>
- Nguyen, T. T. T., & Hamid, M. O. (2016). Language attitudes, identity and L1 maintenance: A qualitative study of Vietnamese ethnic minority students. *System*, 61, 87–97. <https://doi.org/10.1016/j.system.2016.08.003>

- Nyst, V. (2015). The Sign Language Situation in Mali. *Sign Language Studies*, 15(2), 126–150.
- Oriyama, K. (2012). What role can community contact play in heritage language literacy development? Japanese-English bilingual children in Sydney. *Journal of Multilingual and Multicultural Development*, 33(2), 167–186. <http://dx.doi.org/10.1080/01434632.2011.617822>
- Prasad, G. (2012). Multiple Minorities or Culturally and Linguistically Diverse (CLD) Plurilingual Learners? Re-Envisioning Allophone Immigrant Children and Their Inclusion in French-Language Schools in Ontario. *Canadian Modern Language Review*, 68(2), 190–215. <https://doi.org/10.3138/cmlr.68.2.190>
- Pulinx, R., & Avermaet, P. V. (2014). Linguistic diversity and education. Dynamic interactions between language education policies and teachers' beliefs. A qualitative study in secondary schools in Flanders (Belgium). *Revue Française de Linguistique Appliquée*, XIX(2), 9–27.
- Raghavan, P. (2014). It's time to scale the science in the social sciences. *Big Data & Society*. <https://doi.org/10.1177/2053951714532240>
- Riches, C., & Curdt-Christiansen, X. L. (2010). A Tale of Two Montréal Communities: Parents' Perspectives on Their Children's Language and Literacy Development in a Multilingual Context. *The Canadian Modern Language Review / La Revue Canadienne Des Langues Vivantes*, 66(4), 525–555.
- Tang, S. (2010). Foundational Paradigms of Social Sciences. *Philosophy of the Social Sciences*. <https://doi.org/10.1177/0048393109355294>
- Tembe, J., & Norton, B. (2008). Promoting Local Languages in Ugandan Primary Schools: The Community as Stakeholder. *The Canadian Modern Language Review / La Revue Canadienne Des Langues Vivantes*, 65(1), 33–60.
- Toffoli, D. (2015). University students' plurilingual profiles in a French frontier city. *Language Learning in Higher Education*, 5(1), 25–43. <https://doi.org/10.1515/cercles-2015-0002>
- Wei, L. (2011). Multilinguality, Multimodality, and Multicompetence: Code- and Modeswitching by Minority Ethnic Children in Complementary Schools. *The Modern Language Journal*, 95(3), 370–384. <https://doi.org/10.1111/j.1540-4781.2011.01209.x>
- Weston, D. (2011). Gibraltar's position in the Dynamic Model of Postcolonial English. *English World-Wide*, 32(3), 338–367.
- Wisniewski, K. (2017). Empirical Learner Language and the Levels of the Common European Framework of Reference. *Language Learning*, n/a-n/a. <https://doi.org/10.1111/lang.12223>