

# John Benjamins Publishing Company



This is a contribution from *The Politics of Multilingualism. Europeanisation, globalisation and linguistic governance*.

Edited by Peter A. Kraus and François Grin.

© 2018. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible to members (students and staff) only of the author's/s' institute, it is not permitted to post this PDF on the open internet.

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: [www.copyright.com](http://www.copyright.com)).

Please contact [rights@benjamins.nl](mailto:rights@benjamins.nl) or consult our website: [www.benjamins.com](http://www.benjamins.com)

Tables of Contents, abstracts and guidelines are available at [www.benjamins.com](http://www.benjamins.com)

## How to measure linguistic justice?

### Theoretical considerations and the South Tyrol case study of the Calvet Language Barometer

Federico Gobbo

University of Amsterdam / University of Turin

When the concept of linguistic justice was proposed by Pool (1991) in order to cope with the asymmetries quite often found in multilingual contexts, it immediately provoked a great deal of debate. To sum up the debate, there is broad agreement on the meaning of linguistic in-justice, but it is still not clear what linguistic justice really is. This doubt is reflected in the mechanics of the proposed methods for the evaluation of multilingual contexts. What are we measuring? In particular, justice for whom, for example, national citizens or migrants? At which level of analysis, that is to say, at local, national, or transnational? The answers to these types of questions will determine our choice of the right – or at least the appropriate – parameters to be taken into account in order to design an index of the proposed measurement method. In other words, measures are far from neutral, in spite of the fact that they are quantitative, especially in the field of analytical sociolinguistics (Iannàccaro & Dell’Aquila, 2011). After almost a decade of refinements, in particular by Van Parijs (2004, 2011, 2012), Grin (2011), and Grin and Gazzola (2007) observed that the intangible value of a language, being one of the main carriers of culture, is quite often left aside in the indexes that propose to measure linguistic justice as a whole. The only variable that seems to be generally accepted as being relevant is territory.

In this chapter, I examine the Calvet Language Barometer (CLB) in its latest version, published on the web in 2012, under the perspective of linguistic justice. I will proceed backwards, as the CLB aims to measure the “linguistic altitude” of languages in isolation – *i.e.*, their position on the scale – while linguistic justice refers to languages in contact and mainly to multilingual contexts. However, the worldwide gravitational model presented in Calvet (2006, 1999), is naturally compatible with the CLB and can be used for this purpose. The case study of South Tyrol will be presented in order to test the CLB in this perspective, showing to what extent the barometer works, where it does not work, and why. In the conclusion, some preliminary ideas about a genuine multilingual measure of linguistic justice will be presented, based upon an operative notion of multilingual equilibrium.

## 1. Introduction

When dealing with multilingualism, one of the most important methodological issues still left open is how to measure the linguistic context under analysis. Languages live through their speakers: they shape their identity, assess attitudes towards life, and also provide opportunities for mobility – both in terms of levels in society and in terms of movement from one place to another. Intuitively, some languages are in high positions while others not. This is evident in the case of the extremes of the scale: we all agree that English is the language in the highest position nowadays, while endangered languages are in the lowest positions by definition.

However, when we try to clarify what “high” and “low” really means, everything becomes complicated. Grin and Gazzola (2007) already pointed out that quantitative models of the economics of language should take care not only of effectiveness and efficiency (cost-benefit analysis), but also of fairness and communication goals, which are less tangible. In fact, they are influenced by the so-called “symbolic value” nested into languages, quite often left aside in the models that attempt to cope with an ecology of languages – *i.e.*, a social environment in which languages, considered as complex adaptive systems (see Beckner et al., 2009), are in contact. Finding viable parameters for this symbolic value is not straightforward. We find ourselves in a paradox. On the one hand, it is not easy to compare two languages in a given situation by an undisputed set of parameters collected into an index. On the other, we cannot avoid addressing this problem. In fact, every discourse about language diversity presumes a scale of strength – regardless of what this really means.

Gazzola (2014a, 2014b and 2014c) offered an overview of the different perspectives of linguistic justice present in the literature until now, and the need for an index of linguistic justice in order to compare different situations is increasingly evident among the scholars (Alcalde, 2014). Even though linguistic justice was introduced in the 1990s by Pool (Gazzola, 2014b: 4), it was Van Parijs (2004, 2011, 2012) who popularised it not only in academia, but also in the media, focusing on the role of English as a *lingua franca*, a concept highly criticised by many commentators, amongst whom one finds Gazzola and Grin (2013) from a language economy point of view, and the review of Jenkins’ book on the subject by Gobbo (2010), from the perspective of applied linguistics. Defining linguistic justice as an evaluation of the efficiency and fairness of concurrent language-regime scenarios seems to be the most advanced approach in the literature, because it also takes fairness into account (Gazzola & Grin, 2013; Gazzola, 2014c). These studies elaborate on official statistical datasets such as Eurobarometer surveys and Eurostat, which cover broad territories, such as the European Union. Gazzola (2014b) offers a preliminary survey about the two main approaches to linguistic justice, while an extensive overview of the literature can be found in Alcalde (in press).

The first approach is grounded in the theory of justice, based upon liberal egalitarianism, which is the scientific and ideological reference that Van Parijs hails from. Having a minority language as a first language is generally perceived as a handicap – the majority of mankind does not speak the central languages, in Calvet’s terms (see below for details). The second approach praises multilingualism as a value *per se*, as individuals have the human right to develop themselves through their own language, and not the language(s) of others – for example, in schools. In order to have a fair and efficient Europe, for example, this group of authors argues that we need multilingual Europeans – for instance, Marácz and Rosello (2012). The two groups – liberal-egalitarians and “multi-lingualists” – share neither the tools of analysis, nor the results. Indeed, one can conclude that liberal-egalitarians put the individual citizen and his or her rights first, while multi-lingualists put group rights first, and this can partially explain the striking difference in the literature over linguistic hegemony, linguistic (in-) security and linguistic (in-) justice.

From a sociolinguistic point of view, in the official statistical datasets used in the literature, there is no explicit evaluation of the strength of a language (*i.e.*, the prestige from the point of view of the speech community) or the vitality of a language (in the sense of Jules Gilliéron), which can be described as the language loyalty within the speech community, in particular concerning the intergenerational transmission of the language itself (for an analysis of quantitative treatments of language vitality, see Iannàccaro & Dell’Aquila, 2011). Thus, crucial sociolinguistic aspects are simply not taken into account in the literature on linguistic justice, because of the lack of available data and of the different backgrounds of the researchers in the field – not only applied linguistics and translation studies, but also political science and philosophy and economics. For example, key terms such as “planning” or “*laissez-faire* policy” have different meanings, depending on the background of the authors using them (Gazzola, 2014a: 4). This methodological problem is particularly important when we try to depict concurrent scenarios in order to change the *de jure* and *de facto* language policy of the institutional agents in charge. In fact, it is far easier to depict unjust linguistic situations than to propose improvements – if not solutions – in which all the linguistic actors feel they are being treated on an equal basis, in which everybody is in a comfortable position, which we can describe as a position of language security.

At the moment, there is no concrete proposal for measuring linguistic justice through an index, at least according to this author’s knowledge. However, one of the few proposals to capture simultaneously the difference in strength between languages in quantitative and sociolinguistic terms is the Calvet Language Barometer (CLB). Its name comes from the authors, Louis-Jean Calvet and Alain Calvet, respectively a renowned linguist and his brother, a mathematician. In this chapter, the CLB will be analysed as a candidate for measuring linguistic justice. Although

the CLB does not address multilingual settings, but, instead, evaluates every single language, it can be easily applied to situations in which languages are in contact, borrowing the worldwide gravitational model presented in Calvet (2006, 1999). It is important to note that the analysis will not evaluate the CLB *per se*, but only for the purpose of assessing linguistic justice. It is worth noting that, while Calvet (2006, 1999) is well aware of multilingualism and its value *per se*, the CLB is similar to the liberal-egalitarian approach from a methodological point of view. This, the analysis of the CLB can give us useful insights and valid indications in order to build an explicit index of linguistic justice in the near future, thereby overcoming the methodological *impasse* described before. I will proceed backwards. First, the CLB will be presented and discussed. Then, the gravitational model of Calvet (2006, 1999) will be used, in order to introduce a possible use of the CLB in multilingual situations. As Südtirol or South Tyrol is a well-known multilingual case in the sociolinguistic literature, the CLB will be applied to this case study in order to test the CLB itself. The chapter concludes with some general considerations, mainly from a sociolinguistic point of view.

## 2. A barometer to measure the altitude of languages

The CLB was launched in 2010 through a web site,<sup>1</sup> and takes 137 different idioms of the world into account, analysed into ten parameters. In 2012, an update was made, taking 563 languages into account, and an eleventh parameter was added. More parameters can be added in the future, depending on the needs of the users of the CLB itself, according to what is stated on the web site. The authors themselves give some interesting suggestions of possible parameters, such as the number of scientific publications in a given language per year, the production of films in the said language, *etc.* Nonetheless, as there has been no update since 2012, I assume that the CLB is stable.

Clearly, one of the criteria in choosing parameters is completeness. What should be done if data concerning a parameter simply do not exist? The authors suggest assigning a value of zero: for example, setting the number of scientific publications to zero in a minority language belonging to Papua or Cameroon does not radically change the overall picture. Alternatively, it is possible to assign the average value in the column based on similar languages according to the other parameters, so as not to introduce deviations into the analysis. However, it is important to emphasise the fact that these techniques should be avoided whenever possible: they are used

---

1. The actual link is: <http://wikilf.culture.fr/barometre2012>, (last accessed 14 November 2017). No changes have been proposed since 2012.

only for the robustness of the index, as they do not express real-world data. Another important point highlighted by the authors is the fact that there is some overlap between the parameters: in other words, the data are not always independent. For example, it is well known that the Human Development Index is negatively correlated with the total fertility rate. All these limitations should be considered with care when using the CLB in the task of evaluating the degree of linguistic justice in a concrete society.

The weight of the parameters can be adjusted according to the preferences or requirements of the CLB user by means of a slider on the web site. Thus, the CLB is actually a meta-index, because this possibility can generate a number of indexes with very different results, making the CLB both (a) flexible, and (b) not so “user friendly”. Therefore, as a result, all parameters are considered to be equivalent, unless otherwise stated, in order to avoid possible misunderstandings. Each parameter is presented and discussed *per se*, with special intention being given to finding a measure of linguistic justice.

#### Parameter 1: The number of speakers

There is an implicit assumption that the number of speakers is the most important parameter, which is also the parameter generally used by laymen when evaluating the “strength” of a language. In fact, the choice of the languages considered by the barometer is determined by the number of speakers: the baseline used in 2010 was 5 million speakers, while, in 2012, this was reduced to 500,000 speakers. The source of these numbers is the publication *Ethnologue: Languages of the World* (2014). Despite its imperfections – as admitted by the authors on the web site – *Ethnologue* is still the most complete source for this kind of information. On the other hand, the authors argue that *Ethnologue* tends to dissect language varieties too finely, creating a lot of sub-classes, which can cause problems in counting. They cite the case of Malay as an example – in which more than twenty different varieties have been identified. Thus, even the first and most important parameter in the CLB shows evident limits.

The problem of designing the linguistic map, of choosing when to consider a distinctive variety to be a language of its own, is far from straightforward. This problem is well-known in sociolinguistics – as well as in language policy and planning – and it was addressed by Kloss (1967), who introduced the concept of “Ausbau language”. A language is “Ausbau” when it is deliberately elaborated, (re-) shaped, for political and identity reasons, and distinct from another language which is dominant. The general aim is to prevent the Ausbau language from being confused with the other, stronger language in contact. The act of giving a definite

name to the language variety and a distinctive writing system is usually the first step in “Ausbauization” (the term coined by Mauro Tosco), *i.e.*, the process of becoming Ausbau, *i.e.*, how the language is shaped in order to access non-traditional, prestigious domains of use; the first step of the process of becoming Ausbau is to establish a writing norm that is actually used by the speakers. This process cannot be easily measured in absolute terms, as it is put into practice by always keeping the dominant language in mind as a point of reference: the relation between the dominant and the dominated language should be adjusted in order to leave more space to the dominated language, with the intention of maintaining and promoting it. Unfortunately, descriptions of languages usually consider each language as though it were a stand-alone entity, without regard for the multilingual environments in which they live. An example of this can be found in the world map by *Ethnologue* (Lewis et al., 2014), which is an important reference for typological studies.

In this model, the degree of Ausbauization is implicit and, as a result, the proposed taxonomy sometimes produces puzzling results. For example, Greek minorities in Southern Italy, Griko and Grecanic, are not considered to be a variety distinct from standard Modern Greek. In fact, they are all listed with the ISO-code <ell>, instead of having different ISO-codes, which would mark them as distinct languages. In fact, Griko and Grecanic are structurally and sociolinguistically different from Modern Greek: for instance, they are usually written with Latin characters, instead of Greek characters, and – more importantly – they are endangered languages. In fact, the intergenerational transmission of these varieties is not guaranteed: in the Expanded Graded Intergenerational Disruption Scale (EGIDS) – published by Lewis et al. (2010), the same linguists who maintain the publication *Ethnologue: Languages of the World* – these varieties are below level 4, which is the safety line. In contrast, Modern Greek is a language definitely out of danger, being the official language of a well-recognised state, Greece (EGIDS = 1, *i.e.*, Institutional).

The fact that *Ethnologue* is based upon the EGIDS can sometimes be problematic, as the main parameter of the EGIDS to measure linguistic vitality and strength is the number of L1 (first language) speakers and their age: the authors of the CLB cite the case of Swahili, which is spoken as an L2 (second language) much more than as an L1. During the revision of the CLB proposed by Calvet and Calvet in 2012, the new parameter of vehicularity was added, which will be discussed below.

## Parameter 2: Entropy

Entropy is a notion used in physics, and one which has only recently been applied to the field of linguistics. The authors of the CLB refer to the work of Paolillo et al. (2005) which attempted to differentiate between the languages spoken in different

countries from the languages spoken in only one country – in the latter, entropy would be zero. The assumption is that the language spoken in each country is uniform: in other words, there is a perfect match between languages and countries/states. Clearly, it is perfectly understandable that the existence of different countries in which a given language is spoken is a driving force that leads to the emergence of different norms – suffice it to think of the varieties of the Dutch language in Holland and Flanders, for instance. On the other hand, languages often present a lot of variation within the borders of a single country. From a sociolinguistic point of view, the barometer attributes a prominence to the diatopic axis (*i.e.*, the geographic variable in language variation) over the diastratic axis (*i.e.*, the variable that indicates the relation between the social position of the speakers and their language use). In other words, the notion of territory is considered to be the main variable to evaluate the position of languages on the scale, which is a highly questionable assumption, as some languages show a considerable degree of diastratic variation – for instance, German and English (Neumann, 2014). Consequently, languages are evaluated by using the *normative written variety* that they actually show, regardless of their sociolinguistic complexity. Therefore, only languages with a high degree of Ausbau can be considered by the barometer: languages without a stable, written variety cannot be taken into account, regardless of the numbers of speakers who use them. I will come back to this point later.

### Parameter 3: Vehicularity

Although there is a lot of literature about vehicular languages as such, it is rare to find a clear definition of the property of vehicularity in abstract and general terms. The CLB offers a definition in quantitative terms: vehicularity is represented as a ratio between the number of L2 speakers *vis-à-vis* the number of total speakers (L1 + L2). So, a language spoken only as an L1 will have zero vehicularity, while the vehicularity of a language with no L1 speakers will have vehicularity set to one. Interestingly, according to the authors, Hindi is a purely vehicular language, which is hardly believable: in the Indian census data of the last years (1991, 2001, 2011), a significant percentage of Indians indicated Hindi as their mother tongue (data available from the government official web site<sup>2</sup>). Even if the notion of “mother tongue” and L1 are not equivalent (for a critical perspective, see Bonfiglio, 2013, 2010), it seems reasonable to suppose that L1 Hindi speakers *do* exist. Perhaps, the only language which is purely vehicular in this sense is Esperanto, where no monolingual speakers exist: furthermore, even in the case of bilingual families,

---

2. Census Data Online: <http://www.censusindia.gov.in>.



Esperanto has no clear distinctive normative status, compared to non-native, but fluent, language speakers, according to language experts, such as Lindstedt (2010). In any case, the Calvet evaluation of Hindi is determined by the data, which does not touch the definition of vehicularity. In fact, this definition of vehicularity as a ratio is very clear and intuitive, and, in my opinion, it should be used when convenient, even without the whole apparatus of the CLB.

#### Parameter 4: Official status

As acknowledged by the authors, a language can be *de facto* official without holding this status as a result of a legal document: the level of recognition of a language by the political authorities is a complex topic, one in which the level of the authority (local, national, or transnational) should at least be identified properly. The authors use the data provided by Leclerc on his web site entitled *L'aménagement linguistique dans le monde* (language arrangement across the world), which distinguishes between official languages in sovereign states *versus* non-sovereign, attributing a 0.5 value to the latter, with the *caveat* that it counts only once if it is official in different regions of the same sovereign state. However, there are some situations that lead to paradoxes. For instance, the case of Italy is of particular interest here: Italian is official not only in Italy and in Switzerland (both as a federal language and in the Cantons of Ticino and Grisons), but also in the sovereign states of San Marino, the Vatican (with Latin as the state language, which means that it counts 0.5) and the Sovereign Military Hospitaller Order of Saint John of Jerusalem of Rhodes and of Malta. The overall value is 4.5, rank eight in the top ten, after Serbian (value: 4.75) and before Mandarin Chinese (value: 4). It is evident that the political weight of Italy in respect of the status of the Italian language is different from that of San Marino or of the Hospitallers, even though they all value 1. This parameter is particularly problematical, because it does not refer to the speech community in any way, unlike the previous ones. Moreover, there is no consideration of the status of the language in supranational entities, such as the European Union or the United Nations: clearly, being an official language in these institutions contributes to strengthening the language, but this fact is not turned into valuable data.

#### Parameters 5 and 6: The role of translation

Translation is an important means of diffusing ideas which originally belong to another language, so it rightly deserves two parameters, namely, the distinguishing source, and the target languages. Here, the CLB uses the data provided by the *Index Translationum* by Unesco. This index measures the number of translations

by language published since 1979, classified into nine categories: (i) general and bibliography; (ii) philosophy and psychology; (iii) religion and theology; (iv) law, social sciences, and education; (v) natural sciences; (vi) applied sciences; (vii) arts, sports, and games; (viii) literature; (ix) history, geography, and biography. The *Index Translationum* has the advantage of both the coverage of several languages and the neutrality guaranteed by Unesco. However, even this index presents some problems. In particular, the main problem occurs when matching the world map of languages according to *Ethnologue* with the world map of Unesco. The authors are aware of this. The case of Serbo-Croatian of the former Yugoslavia, now split up into Serbian, Croatian, Bosnian and Montenegrin, as well as the complex situation of Arabic, are discussed by the authors. Essentially, they re-framed the data according to the ISO codes of the languages, which is a reasonable solution – but the problem still remains. From a methodological point of view, the CLB is calculated here by using indirect data, *i.e.*, raw data that were already collected for other purposes and transformed independently. In other words, these two parameters devoted to translation are not based upon data, but upon another index instead, so that the CLB becomes partially a meta-index. The same is true for parameter 9, as we will see below.

#### Parameter 7: International literary awards

Literature is a factor that testifies to the high level of the Ausbau of a language: only a small percentage of the languages of the world are recognised as vehicles of world literature through prizes and awards. This parameter has the previous ones as its basis. In fact, works must be known in an international setting before they are nominated for an international literary award, such as the Nobel Prize for Literature. There are a number of reasons why the Nobel Prize cannot suffice as the only source for this parameter: in particular, the eurocentrism and the left-wing orientation of the political orientation of the Nobel committee. Nominees count 0.5 (only once, if nominated many times), while winners count 1. Nominees and winners are considered according to the languages in which they write, not their nationality. For example, Ngũgĩ wa Thiong'o is counted twice, for English and for Gikuyu – which is admittedly questionable if we think of the reasons that led him to use English (Thiong'o, 1986). The authors are aware that the parameter is imperfect, but it is nevertheless important for the self-esteem of the speakers of the languages, where international literary awards exist. Perhaps, it would be better to consider it as a correction or control factor of the *Index Translationum* taken globally. However, this is, admittedly, possible if the user of the CLB changes the weight of the different parameters.

### Parameter 8: The number of Wikipedia articles

This parameter considers the “grand total” of the articles on Wikipedia by September 2011. This parameter shows severe limits, due to the internal policy of Wikipedia. First, it does not consider real articles differently from stubs, which can even be generated by software. There is already a specific episode about Wikipedia in Volapük, a language planned for international purposes before Esperanto, whose interest now is merely historical: more than 100,000 stubs were generated by a single user with the help of a software robot, and, after a discussion in January 2008, Wikipedia administrators decided to keep the articles anyway.<sup>3</sup> This was a very extreme case, which abruptly pushed Volapük into the second range of languages (more than 100,000 articles) with only 30 active users registered.<sup>4</sup> To make a comparison, Armenian has a similar number of articles, but it has 209 active users. Another problem is how to count languages with more than one writing system, such as Norwegian, which has two distinct Wikipedias. I think that the number of active users should be taken more seriously into account in analysing the impact of Wikipedia. For example, if we list the top ten Wikipedias, according to the number of active users, Dutch is no longer in second place, but goes down to eighth place: this means that Dutch language contributors, who number more than 4,000, are very active in opening new pages, compared, for example, to the Polish language, which has a similar number of contributors, but few articles (+1,700,000 for Dutch, a bit more than one million for Polish). Moreover, Hale (2014) pointed out that the role of multilingual users (just over 15%) is important in the maintenance and growth of Wikipedias. In fact, multilingual users are far more active than their monolingual counterparts, and act as concrete bridges between the different versions. In this specific case, an index of the language vitality of Wikipedia could be made upon the basis of fine-grained data, as a lot of data suitable for statistical analysis are always available and are updated monthly or even daily in some cases. At the very least, the number of active speakers compared to the number of the total speakers (L1 + L2) could be performed, in order to form an index of the activity of the various Wikipedias. The simple “grand total” of Wikipedia articles is not enough.

---

3. [http://meta.wikimedia.org/wiki/Proposals\\_for\\_closing\\_projects/Radical\\_cleanup\\_of\\_Volap%C3%BCK\\_Wikipedia](http://meta.wikimedia.org/wiki/Proposals_for_closing_projects/Radical_cleanup_of_Volap%C3%BCK_Wikipedia).

4. [http://meta.wikimedia.org/wiki/List\\_of\\_Wikipedias#Grand\\_Total](http://meta.wikimedia.org/wiki/List_of_Wikipedias#Grand_Total) (last accessed 16 June 2014).

### Parameters 9 and 10: Human Development Index (HDI) and total fertility rate

If, in the case of translation, an index was a source of data, in this case, an index is used as it is. The Human Development Index (HDI) is used by the UN to measure the achievements of countries in education, health, and income. It is calculated for the member countries of the UN, not for their languages. However, the explanation by Calvet and Calvet here is insufficient: after all, the HDI and the total fertility rate are indexes unrelated to languages. In fact, surprising results are obtained: classification solely according to this criterion gives different results when it considers languages belonging to the same country. For instance, the Hawai'i Creole English is ranked in second position while English (American? English? Global?) is in fourth position: although the distance between the two is small (the difference is only 0.021), it is not clear *why* they are different. I think that the use of this index for languages is spurious, and should not be used at all. Similar issues arise for the total fertility rate, which is part of the report in which the HDI is also used, and it is again calculated for countries, not languages. At most, Parameters 9 and 10 can be allowed, but with very little weight.

### Parameter 11: Language use in the Internet

This parameter can be put into relation with Parameter 8 concerning Wikipedia, as both illustrate the presence of the language on the web. However, if the data offered for Wikipedia by its Foundation are reliable, I doubt that these Internet World Stats really are. In fact, the source of the web site<sup>5</sup> is the Miniwatts Marketing Group, a limited liability company, legally established in 1997 in Bogotá, Colombia. Furthermore, the data here are already an elaboration by the Miniwatts company based upon the US Census Bureau, Nielsen Online, and other sources, not all explicitly listed: it is simply impossible to double-check the data. It is very different from a public institution such as the UN, the source of the two previous parameters, or the Wikipedia Foundation, where raw data are always available, because of the free software licence. As in the case of Parameters 9 and 10, the influence of Parameter 11 should not be kept to the minimum.

---

5. <http://www.internetworldstats.com/stats.htm>, last accessed 16 June 2014.

Two versions of the barometer: Unweighted *versus* weighted

In Table 7.1 (below), the reader can see a comparison of the original CLB and with different weights for different parameters, after their discussion. In particular, the weight of the most problematic parameters were given a weight of 0.25 (international literary awards, HDI and total fertility rate, and language use on the Internet), while the less problematic parameters were given a weight of 0.5 (official status, and translation). The other parameters that do not present particular problems were left as they are.

**Table 7.1** An application of the Calvet Language Barometer with and without weights

Parameters			Weight		Weight	
Number of speakers			1		1	
Entropy			1		1	
Vehicularity			1		1	
Official status			1		0.5	
Role of translation			1		0.5	
International literary awards			1		0.25	
Number of Wikipedia articles			1		1	
Human Development Index			1		0.25	
Total fertility rate			1		0.25	
Language use in the Internet			1		0.25	

Rank	Language	Score	Rank	Language	Score
1	English	9.062	1	English	5.427
2	Spanish	7.806	2	Spanish	5.131
3	French	7.733	3	French	4.580
4	German	6.987	4	German	4.108
5	Russian	6.335	5	Russian	4.024
6	Japanese	6.187	6	Mandarin Chinese	3.918
7	Dutch	6.138	7	Portuguese	3.792
8	Italian	6.131	8	Italian	3.684
9	Portuguese	5.97	9	Dutch	3.670
10	Mandarin Chinese	5.964	10	Japanese	3.635
11	Swedish	5.543	11	Hindi	3.382
12	Turkish	5.321	12	Turkish	3.309
13	Norwegian	5.232	13	Bengali	3.288
14	Polish	5.2	14	Rumanian	3.231
15	Danish	5.104	15	Farsi	3.157

Source: Elaboration from the CBL web site (data calculated 19 June 2014).

The top five languages are left untouched by the change: English, Spanish, French, German, and Russian are the highest languages in the barometer. On the other hand, there is a significant change in Mandarin Chinese, which moves up from position 10 to position 6, switching positions with Japanese. The re-organisation of the influence of the parameters also lets some non-Western languages climb up the scale: Hindi, Bengali and Farsi show up in the first fifteen, while the three main Scandinavian languages (Swedish, Norwegian, and Danish) disappear from this group.

The CLB has the merit of being very flexible: the weighted model, however, which takes the observations made until now into account, presents a view of the languages of the world that is closer to reality. On the other hand, the CLB is not robust enough to be used as such as an index of linguistic justice.

### 3. The gravitational model and the barometer

Before the launch of the CLB in 2010, Louis-Jean Calvet had worked for years on the ecology of world languages (1999, 2006). Several authors had used the word “ecology” in different ways; Calvet considers languages as “ecolinguistic systems” (2006: 46), using ecology in the sense of Immanuel Wallerstein’s World systems theory. Systemics, as mentioned above, are specifically applied to linguistics by Beckner et al. (2009) in considering languages as complex adaptive systems. In Calvet’s view, each language is part of an ecological “niche”, constituted by its relations with other languages, “by the place it occupies in the ecosystem, *i.e.*, by its functions and by its relations with the environment – essentially, that is, by geography, which plays a defining role in the spread of languages” (2006: 24). It is clear that the defining role of the ecolinguistic system is – again – territory, which drives the whole analysis. If we want to apply the CLB to measure linguistic justice, we have to identify a definite ecolinguistic system as the context under scrutiny. However, as we have already seen, the CLB cannot be directly applied to any ecolinguistic system, because languages are placed in the CLB according to the absolute number of L1 speakers worldwide, regardless of the variable of territory. Only in some cases are languages located in specific areas, in particular, minority languages and Creole languages. For instance, minority languages such as Sicilian, Lombard and Piedmontese are mainly rooted in specific regions of Italy – although the presence of communities abroad has its importance, mainly for symbolic reasons.

In order to try to solve the problems illustrated so far, let us take the gravitational model of Calvet (2006, 1999) to compare the weight of different languages. Unlike the CLB, this model is qualitative. The gravitational model was originally

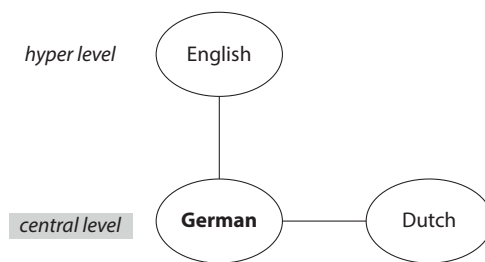
introduced by Abram de Swaan (2001), and it is still in use – for example, by Hiddinga and Crasborn (2011) for sign languages. The basic idea comes from astrophysics: the world is seen as a galaxy in which languages are the stars. If a star has a higher mass, it will be in a more prominent position. In the terms of the CLB, it will be in a higher position on the scale. Following this metaphor, bilingual speakers form a constellation of languages, as they are the connectors. Some languages are more attractive than others, and a dynamic equilibrium between central and peripheral languages emerges, with four levels of weight in total.

The reference-point of the galaxy is currently English, the hyper-central language: L1 speakers tend to be monolingual, while non-natives tend to learn it as an L2. This situation is called *vertical bilingualism*: an L1 speaker of a peripheral language shows a tendency to learn a more central language. This can be reflected in the CLB in terms of vehicularity and the *Index Translationum*: both should show high values. At the second level, there are the super-central languages: Arabic, Chinese, French, Hindi, Malay, Portuguese, Russian, Spanish, and Swahili. Calvet observes that L1 speakers tend either to be monolingual or vertically bilingual with English. In terms of the CLB, all these languages are official (Parameter 4) in at least one sovereign state, and they also have good scores in terms of vehicularity and in the *Index Translationum*. At the third level, there are more or less one hundred central languages, where L1 speakers show a tendency towards vertical bilingualism with super-central languages, while the 5,000 peripheral languages (level 4) show not only vertical, but also horizontal, bilingualism – when a speaker acquires an L2 of the same weight (it should be in a similar position in the CLB). According to the author of the galaxy model, Jean-Louis Calvet, vertical bilingualism proceeds step-by-step: in Senegal, a speaker of Serere or Diola (level 4) will acquire Wolof (level 3) before French (level 2), and finally English (level 1; Calvet, 2006: 61). Henceforth, in the case of the very complex ecolinguistic systems in Africa, four levels of analysis are needed.

However, it seems to me that several educational policies in various parts of the world tend to jump directly to learn level 2 languages (such as French) and, in particular, the level 1 language, to wit, English, not following the multiple-stage process illustrated above. Moreover, in many ecolinguistic systems – for instance, in Europe or in Canada – it is sufficient to have three levels in total, instead of four. This raises the question of the appropriateness of the names of the classes: in particular, I doubt that L1 speakers of Occitan or Piedmontese ever refer to their own languages as “central”, following Calvet’s proposal, especially if we take into account the fact that there is no periphery (*i.e.*, level 4) left.

There is another limitation in Calvet’s gravitational system, which is not linked to a particular ecolinguistic systems, but turns out to be a theoretical one. The

model assumes that every language is connected with one – and only one – other language – like planets and satellites. In several contexts, the relations between different languages residing in the same ecolinguistic niche are more complex than this. Let us consider L2 languages learnt because of mobility or motility (Houtkamp, 2014) – cases which are not explicitly dealt with by the author. For instance, a bilingual German-English person working and living in The Netherlands would probably learn Dutch in order not to be excluded from important domains of the life of the inhabitants of the ecolinguistic system in which he or she happens to live. For this reason, I argue that typed connections – in the sense used in graph theory – between languages should be introduced here: in other words, there is more than one possible link between languages. Calvet (2006: 61) distinguishes between spontaneous learning and programmed learning (for example, at school) in the case of vertical and horizontal bilingualism. While verticality and horizontality are easily depicted by the directions of the connections, I will use dotted lines for spontaneous acquisition (informal learning), while continuous lines will be used for programmed, structured learning (formal and non-formal learning). Programmed learning does not block spontaneous language acquisition, rather, it reinforces it, and, for this reason, a stronger, continuous line was chosen; in contrast, spontaneous learning is mainly context-dependent, and, in general, is weaker. Furthermore, the presence of an explicit language-learning policy changes the symbolic value of the L2 language: under a *ceteris paribus* assumption, a programmed learning strategy is stronger than a spontaneous one. Finally, the symbolic relevance of an L1 should be depicted accordingly: therefore, the names referring to L1 will be put in bold type. Figure 7.1 below illustrates the example presented above.



**Figure 7.1** A complex ecolinguistic system with both vertical and horizontal bilingualism

In sum, it seems more useful to apply *some* parameters of the CLB, instead of the whole apparatus, according to the case under analysis. In fact, a majority language in absolute terms can become a minority language in a specific ecolinguistic system. Each parameter should be selected carefully, with a proper weight, possibly avoiding



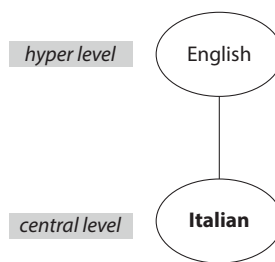
the most problematical ones – especially those based upon the existing indexes. Then, Calvet’s definition of vertical *vs.* horizontal bilingualism can be applied on a three or four-level scale, again, according to the context. Roughly speaking, African contexts need four, while European ones seems to need only three in most cases. Typical links for bilingualism are proposed here, on the basis of the distinction between spontaneous language acquisition *vs.* programmed language learning. The symbolic value of the mother tongue – with an oversimplification, corresponding to the L1 – is retained as the starting-point of the multilingual situation, finally permitting a description of the strength of the specific language in the context, and the attitudes towards language learning by the prototype actors, defined along with their mother tongues.

I argue that a description of the languages present in a specific ecolinguistic system is pre-condition to any consideration about linguistic justice. The final goal is to find a non-unjust language policy for every actor, so that, on the one hand, the language learning effort would be more or less the same for everybody, while, on the other, the actors’ communicative possibilities would be the highest. My point is that a sociolinguistic notion of linguistic justice, which we can call *sociolinguistic justice* here, should be investigated in depth, in which the learning efforts and communicative possibilities are the main concurring variables, so that the context could be described as a dynamic (socio-) linguistic equilibrium. Now, after these necessary adaptations, we can apply this set of tools to a case study.

#### **4. An application: The barometer of the linguistic justice in South Tyrol**

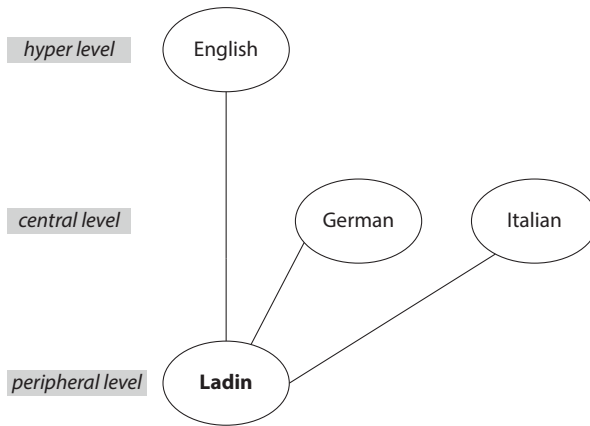
South Tyrol is often considered a success case of language management. The ecolinguistic system of this area, which is politically part of Italy, with a complex historical background linked to the Austrian Empire, and with a geographic border with Switzerland, contains three distinctive language groups: Germans, Italians and Ladins. Traditionally, the overall political strategy to prevent conflicts can be described as “separation under the same roof”: the three groups should avoid contact as much as possible, also thanks to the fact that the territory is very mountainous, while formal co-operation was guaranteed through the institutions which are, to a large extent, autonomous from the central government in Rome. According to Pallaver (2014), there is a movement towards a new strategy of an associative conflict-resolution model, in which co-operation between the three linguistic groups operates at different levels, starting from initiatives coming from civil society.

The multilingual landscape of South Tyrol is therefore complex, as there is clear evidence of the contacts between the three languages over a long period of time. We will refer to the results of the four-year research project *Kontatto*, which ended in May 2014, and which aimed to study in detail what happens in terms of identity in a border territory like South Tyrol. The endogenous linguistic systems there belong to different groups: standard German and the local German dialect are Germanic, while Italian and Ladin are Romance languages. Their constant and close contacts for decades (if not centuries) have led to the phenomena of linguistic innovation, induced by contact itself. Italian is spoken mainly in the cities of Bozen-Bolzano and Meran-Merano as well as in the Bassa Atesina area, while German is spoken by the majority of the population, mainly in its local, non-standard Germanic dialects. On the other hand, Ladins are located in the valleys where they belong by tradition, in particular, Badia and Gardena, if we consider South Tyrol alone as our ecolinguistic “niche”. The degree of multilingualism can vary a lot, considering that South Tyroleans can have three different L1s: in particular, Italians in the cities tend to be monolingual, while Germans shows a tendency towards the programmed acquisition of Italian: finally, Ladins often acquire both. In the Bassa Atesina, the situation is more fluid, but, in any case, it is unlikely that Ladin will be acquired by the L1 speakers of Italian and German (for research in this field, see Meluzzi et al., 2013). The three prototypical situations, according to the different mother tongues considered, are illustrated in Figures 7.2 and 7.3 below. However, English always plays a hyper-central role, which is also made institutionally clear in the language policy of the Free University of Bozen/Bolzano, which is officially trilingual (English, German, and Italian) and intercultural, according to its motto.<sup>6</sup>



**Figure 7.2** The ecolinguistic system of South Tyrol from the perspective of a prototypical L1 Italian

6. See the web site: <http://www.unibz.it>, last accessed 27 August 2014.



**Figure 7.3** The ecolinguistic system of South Tyrol from the perspective of a prototypical L1 Ladin

From a sociolinguistic point of view, the three prototypical L1 speakers are too simple. On the other hand, our aim is to compare the efforts of the typical multilingual strategies and the attitudes of the three different communities in order to identify linguistic injustice and to overcome it whenever possible. It is clear that the Ladins are keen to acquire a richer multilingualism than the others: their language effort is higher. Thus, while obtaining a communicative power similar to German L1 speakers, they are clearly in an unjust position. On the other hand, the vertical multilingualism of the Italian L1 speakers is scarce: their language effort is low compared to that of the others, and thus they obtain a slightly lower communicative power, compared to the other two groups.

Unfortunately, given that it is a lesser-used language, Ladin is not considered in the CLB, and thus a direct comparison between the languages using the other parameters is simply impossible. However, as stated above, the most fruitful way to use the CLB is to look at each single parameter, according to the concrete needs of comparison. For example, in this case, the official status (parameter 4) clearly confirms the disadvantages of Ladin, compared to German and Italian. A helpful parameter should be calculated on the basis of the *Index Translationum*, but it is limited to the territory under analysis, as well as a calculation of the vehicularity rate within South Tyrol, in order to put the situation in equilibrium.

## 5. Final observations

It is very difficult to calculate the weight of any language, although the CLB is a concrete proposal in this direction. Even if severe limitations have been found, some ideas contained within it can be used as a first step towards establishing a comprehensive methodology to evaluate the linguistic (in-) justice of a given ecolinguistic system, identified through its territory and through more robust sociolinguistic data. The general aim is to find the optimal equilibrium between the following two related variables: first, the language-learning effort, and then the corresponding communicative power obtained. The fine-tuning of this interpretation of linguistic justice as balanced multilingualism remains a task for the future.

## References

- Alcalde, Javier (in press). “Linguistic Justice: An Interdisciplinary Overview of the Literature”, in: Michele Gazzola, et al (eds), *Language Policy and Linguistic Justice: Economic, Philosophical and Sociolinguistic Approaches*. Berlin: Springer Verlag. Pre-print published online at: <http://multilingualism.humanities.uva.nl/download>.
- Alcalde, Javier (2014). “On the Need for a Linguistic Justice Index”, *Peace in Progress: Linguistic Conflicts*, No. 21.
- Beckner, Clay, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman and Tom Schoenemann (2009). “Language is a Complex Adaptive System: Position Paper”, *Language Learning*, 59 (Supp. 1): 1–26. <https://doi.org/10.1111/j.1467-9922.2009.00534.x>
- Bonfiglio, Thomas Paul (2013). “Inventing the Native Speaker”, *Critical Multilingualism Studies*, 1 (2): 29–58.
- Bonfiglio, Thomas Paul (2010). *Mother Tongues and Nations: The Invention of the Native Speaker*. New York: Walter de Gruyter. <https://doi.org/10.1515/9781934078266>
- Calvet, Louis-Jean (1999). *Pour une ecologie des langues du monde*. Paris: Plon.
- Calvet, Louis-Jean (2006). *Towards and Ecology of World Languages*. Cambridge: Polity.
- Calvet, Louis-Jean, and Calvet Alain (2012). “Baromètre Calvet des langues du monde. WikiLF”. Published online at: <http://wikilf.culture.fr/barometre2012>.
- Gazzola, Michele (2014a). “Partecipazione, esclusione linguistica e traduzione: Una valutazione del regime linguistico dell’Unione Europea.” (Participation, Linguistic Disenfranchisement and Translation: An Evaluation of the Language Regime of the European Union). ELF Working Paper: #12. Université de Genève.
- Gazzola, Michele (2014b). “Lingva justeco: kiel taksi ĝin? La ekzemplo de Eŭropa Unio.” (Linguistic Justice: How to Evaluate it? The Example of the European Union), ELF Working Paper: #14. Université de Genève.
- Gazzola, Michele (2014c). “Language Policy and Linguistic Justice in the European Union: The Socio-Economic Effects of Multilingualism.” ELF Working Paper: #15. Université de Genève.

- Gazzola, Michele, and François Grin (2013). "Is ELF more Effective and Fair than Translation? An Evaluation of the EU's Multilingual Regime", *International Journal of Applied Linguistics*, 23 (1): 93–107. <https://doi.org/10.1111/ijal.12014>
- Gobbo, Federico (2010). "Review of: Jennifer Jenkins. English as a Lingua Franca: Attitude and Identity", *Language Problems & Language Planning*, 34 (1): 81–83. <https://doi.org/10.1075/lplp.34.1.05gob>
- Grin, François (2011). "Using Territoriality to Support Genuine Linguistic Diversity, not to get Rid of it", in: Philippe Van Parijs (ed.), *The Linguistic Territoriality Principle: Right Violation or Parity of Esteem?*, pp. 28–33. Brussels: Re-Bel, 11, available at: [www.rethinkingbelgium.eu](http://www.rethinkingbelgium.eu).
- Grin, François, and Michele Gazzola (2007). "Assessing Efficiency and Fairness in Multilingual Communication: Towards a General Analytical Framework", *AILA Review*, 20, pp. 87–105.
- Hale, Scott A. (2014). "Multilinguals and Wikipedia Editing", *WebSci '14*. <https://doi.org/10.1145/2615569.2615684>.
- Hiddinga, Anja, and Onno Crasborn (2011). "Signed Languages and Globalization", *Language in Society*, 40 (4): 483–505. <https://doi.org/10.1017/S0047404511000480>
- Houtkamp, Christopher (2014). "Integrating Language in Theories on Long-distance Movement: Migration vs Mobility & the Concept of Mobility". *A'dam Multiling*. Working Paper #2.
- Iannàccaro, Gabriele, and Vittorio Dell'Aquila (2011). "Numeri soggettivi. Spunti sulla vitalità linguistica da inchieste e trattamenti quantitativi", in: Bruno Moretti, Elena Maria Pandolfi and Matteo Casoni (eds), *Vitalità di una lingua minoritaria. Aspetti e proposte metodologiche* (The Vitality of a Minority Language. Aspects and Methodological Issues), pp. 151–192. Bellinzona: Osservatorio Linguistico della Svizzera Italiana.
- Kloss, Heinz (1967). "Abstand Languages and Ausbau Languages", *Anthropological Linguistics*, 9 (7): 29–41.
- Lewis, M. Paul, Gary F. Simons and Charles D. Fennig (eds) (2014). *Ethnologue: Languages of the World, Seventeenth Edition*, Dallas TX: SIL International, available at: <http://www.ethnologue.com>, accessed repeatedly since its publication.
- Lewis, M. Paul, and Gary F. Simons (2010). "Assessing Endangerment: Expanding Fishman's GIDS", *Revue Roumaine de Linguistique*, 55 (2): 103–120.
- Lindstedt, Jouko (2010). "Esperanto as a Family Language", in: Fred Dervin (ed.), *Lingua francas. La véhicularité linguistique pour vivre, travailler et étudier*, pp. 69–80. Paris: L'Harmattan.
- Marác, László, and Mireille Rosello (eds) (2012). *Multilingual Europe, Multilingual Europeans*. Amsterdam-New York: Rodopi. <https://doi.org/10.1163/9789401208031>
- Meluzzi, Chiara, Simone Ciccolone and Iliaria Fiorentini (2013). "Contact-induced Innovation in a Multilingual Setting: Evidence from Italian, German, and Ladin in South Tyrol", 9th UK Language Variation and Change Conference (UKLVC9), Sheffield.
- Neumann, Stella (2014). *Contrastive Register Variation: A Quantitative Approach to the Comparison of English and German*. Berlin: Walter de Gruyter.
- Paolillo, John, Daniel Pimienta and Daniel Prado, et al (2005). *Mesurer la diversité linguistique sur Internet* (Measuring Linguistic Diversity on the Internet). Paris: UNESCO.
- Pallaver, Günther (2014). "South Tyrol's Changing Political System: From Dissociative on the Road to Associative Conflict Resolution", *Nationalities Papers: The Journal of Nationalism and Ethnicity*, 42 (3): 376–398. <https://doi.org/10.1080/00905992.2013.856393>
- Pool, Jonathan (1991). "The Official Language Problem", *American Political Science Review*, 25 (2): 485–514.
- Swaan, Abram de (2001). *Words of the World: The Global Language System*. Cambridge: Polity Press.

- Thiong'o, Ngũgĩ wa (1986). *Decolonising the Mind: The Politics of Language in African Literature*. Nairobi: East African Publishers.
- Van Parijs, Philippe (2004). "Europe's Linguistic Challenge", *European Journal of Sociology*, 45 (1): 111–152. <https://doi.org/10.1017/S0003975604001407>
- Van Parijs, Philippe (2011). *Linguistic Justice for Europe and for the World*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:osobl/9780199208876.001.0001>
- Van Parijs, Philippe (2012). "On Linguistic Territoriality and Belgium's Linguistic Future", in: Patricia Popelier, Dave Sinardet, Jan Velaers and Bea Cantillon (eds), *Belgium: Quo Vadis? Waarheen na de zesde staatshervorming*, pp. 35–60. Antwerp: Intersentia.

