Introduction

Wendy Ayres-Bennett and John Bellamy

1   Introduction

This handbook of language standardization presents both a state of the art of current work on standardization and cutting-edge research on the subject. It offers an original contribution to the important field of 'comparative standardology' (Joseph 1987: 13), bringing together a wide range of case studies, including those which examine 'classic' examples of the standardization of European national languages and others which challenge accepted wisdom and traditional accounts. The breadth of languages and situations treated opens up new perspectives on language standardization and the opportunities and challenges it offers to speakers, communities and nations. Five parts focus in turn on models and theories of standardization (Part I), questions of authority and legitimacy (Part II), literacy and education (Part III), borders and boundaries (Part IV) and standardization in Late Modernity (Part V).

Linguistic standardization has long preoccupied researchers working in different sub-disciplines of linguistics, including historical linguistics, applied linguistics and sociolinguists, as well as those working on language policy. It has equally been studied from a number of different approaches, which will be discussed in more detail below (as well as by Ayres-Bennett in Chapter 1). Since the 1950s, an important body of work has focused on the processes – and the potentially related stages – of standardization. Much of the work in historical linguistics and language policy starts, for instance, by referring to the seminal work of Einar Haugen and the processes of selection, elaboration, codification and implementation of the standard (Haugen 1966). More recently, other processes, such as focusing,

micro-accommodation, levelling and dialect convergence, have been added to the discussion (cf. Deumert 2004: 2–3). Early on, typologies of standard languages also emerged (e.g. Kloss 1968), embracing issues such as the type of country, the language's developmental status, the judicial status of the speech community and the language's numerical strength. Subsequent typologies, such as that proposed by Auer (2005), which examines how dialect/standard constellations change over time, introduce a chronological dimension into the discussion. Other studies focus rather on the historical, political and social factors – whether local or global – which favour standardization, such as urbanization and industrialization, nationalism and nation-building, the spread of education and the role of the mass media in diffusing the standard. More recently, the effects of globalization both on standardization and linguistic diversity have come to the fore, leading to questions as to the extent to which standardization promotes or harms the survival of endangered or minoritized languages (e.g. Lane et al. 2017). Some scholars, such as Del Valle (2013), favour a glottopolitical approach (Guespin & Marcellesi 1986) that attempts to uncover the political foundations that underlie the debates and controversies.

The rich diversity of approaches to linguistic standardization is at once stimulating and challenging. In the next section, we turn to some of the key issues that run through the literature on standardization as well as the chapters in this volume, and on which there has not always been a clear consensus. We start with terminological issues (Section 2.1), before considering various recurrent issues regarding the nature of the standard (Section 2.2). In Section 3, we discuss new material and data that are leading scholars to review and rewrite traditional accounts of standardization, including multilingual contexts (Section 3.1) and minoritized, transnational and transcultural situations (Section 3.2). We then present the contribution of historical sociolinguistics to the elaboration of new approaches (Section 3.3), before considering how the primacy of the written word in standardization is being challenged (Section 3.4). Section 4 is concerned with two important areas in which standard languages have real-world impacts: language rights and education. We conclude with a brief discussion of the changing role and nature of standard languages in Late Modernity (Section 5).

2   Key Recurrent Issues

2.1   Terminological Issues

A thorny issue in studies of language standardization is the question of terminology, which is not yet consistently applied across different studies. Variation in usage concerns even the most basic terms such as 'standard' and 'norm'. Whilst some scholars use 'standard' and 'norm' quasi-interchangeably, others distinguish them: for Bartsch (1987: 248), for instance,

the concept of 'norm' is not restricted to the standard variety, and so the expression 'norms of the standard variety' is possible (Bartsch 1987: 248). Differences also affect the subcategorization and analysis of these terms. Some analyses rely on making a simple dichotomy between descriptive and prescriptive norms (e.g. Rey 1972; McGroarty 1996: 22–3; Auroux 1998: 240–3), whilst others elaborate more types of norm, such as mandatory, permissive and power-conferring norms in the case of Bartsch (1987: 80). There is also variation across different language traditions and different chronological periods.[1] In Chapter 11, Maraschio and Matarrese cite the case of Italian, where a whole host of competing terms have been used to refer to the standard, including 'common Italian', 'good Italian', 'literary Italian', 'classic Italian', 'correct Italian', 'Italian without adjectives', 'supraregional Italian', 'normal Italian' and 'normed Italian'. The potential overlap or fuzzy boundaries between these different labels is obvious. Another example is the difference between 'non-standard' and the more pejorative 'substandard', both attested in their linguistic usage in English for the first time in the first third of the twentieth century according to the Oxford English Dictionary (OED), but still often not kept clearly apart.

That is not to say, of course, that certain terminology has not become established. For instance, the distinction made by Joseph (1987) between 'language standards' and 'standard languages' is now broadly accepted, and although the terms 'polycentric' and 'compositional' standards are still sometimes used, since 1992 most scholars have followed Clyne (1992) in using the expression 'pluricentric language' to refer to a language which has different interacting language standards that are often associated with different countries (e.g. Swiss German, Austrian German, alongside the standard German of Germany). Throughout the volume, there is also, for instance, recurring usage of Haugen's categories of selection, elaboration, codification and implementation (discussed by Ayres-Bennett in Chapter 1) and acceptance of a distinction between standardization 'from above' (Rutten & Vosters, Chapter 2) and 'from below' (Elspaß, Chapter 3).

Another complex question is the relationship of the standard with two other sets of terminology, which refer to the status of different languages. The first of these is the ten-point Expanded Graded Intergenerational Disruption Scale,[2] which relates to the overall development versus endangerment of a language, ranging from International and National through to Dormant or Extinct (cf. Fishman 1991). The second relates to the degree of official recognition of a language, ranging from Statutory national languages through to Provincially recognized languages or Languages of recognized nationality. Whilst in the nineteenth century standard language, national language and official language were all broadly equated, as the ideologies of language and nation went hand in hand (see Rutten & Vosters,

---

[1] An example of this is the term 'koine', discussed in Chapter 1.

[2] www-ethnologue-com.ezp.lib.cam.ac.uk/about/language-status

Chapter 2), this simple equation is untenable for many languages of the world. In Luxembourg, for example, an officially multilingual nation, the Language Law of 24 February 1984 established Luxembourgish as the country's 'national' language (Article 1), whilst French was granted the status of legislative language (Article 2) and French, German or Luxembourgish were designated for all administrative matters (Article 3).[3] The case of Catalan (discussed by Nadal & Feliu in Chapter 22), a Statutory provincial language, raises important questions around regional autonomy and territorialism. We will return to the question of standardization for nationhood and autonomy as well as for minoritized languages below in Section 3.

2.2   Questions about the Nature of the Standard

2.2.1   Written Standards and Spoken Standards

Traditional views of a standard language have primarily conceptualized it as a written form of the language. Although Ferguson contends in his classic study of diglossia that the L variety or the local dialect(s) of the language 'may include a standard or regional standards' (Ferguson 1959: 336), his description of the H variety links it mainly with formal writing. Haugen (1972: 246) later reinforces this association by observing that 'it is a significant and probably crucial requirement for a standard language that it be written', and in later versions of his standardization model, he specifically includes the *Graphization* stage (Haugen 1987: 64), where a written code is developed and elaborated, particularly for unwritten varieties.

The primacy of writing for developing language norms is also recognized by Auroux (1992: 28–9) and continues in work by Milroy (2001: 525), who considers writing to be characteristic of a predominantly 'top-down' process of imposing a standard language ideology (discussed in more detail in Section 2.2.4), for which correctness is a fundamental part of nurturing a standard. The boundary marking which arises from preserving a standard (i.e. deciding which linguistic features are standard and which are not) leads to a heightened sense of correctness, and the written form is easier to control and conserve than the spoken. Being less dependent on the abstraction of writing, spoken language has inherently more variability and comparatively is less often subject to regulation. To use the example of Irish, Ó Murchadha (2013: 79–81) discusses the tensions surrounding Gaeltacht and post-Gaeltacht spoken norms, although the written form, *An Caighdeán Oifigiúil*, is a unitary standard written variety.

It is mainly the standard written language which is used for literature, laws and formal institutional documentation which imbues the written form with prestige and often also makes it emblematic of the nation-state. The standard language is commonly linked with nationally renowned

[3] http://legilux.public.lu/eli/etat/leg/loi/1984/02/24/n1/jo

writers and poets, such as Cervantes for Spanish and Goethe for German, and celebrated literary works, such as *Blanquerna* (ca. 1283) for Catalan, further reinforcing the connection between a standard and a respected written form (see also Maraschio & Matarrese, Chapter 11). Popular conceptions of standard language associate it with dictionaries and grammars, which are authoritative published works that language users turn to for guidance on norms and 'proper' usage (cf. Ammon's (2003: 59) notion of the *language codex*).

Therefore, the most highly regulated domain is the spelling system because only in the orthography can standardization be fully realized (Milroy & Milroy 2012: 18). Spelling is especially amenable to control since it is relatively easy to oversee, it is unambiguous and it is straightforward to correct. Deviations from orthographical norms are rarely tolerated in formal contexts, although this is not the case for writing practices in the private domain. Sebba (2007: 5), citing Strang (1970: 170), notes that 'even though the norms of English spelling became highly standardized during the seventeenth century, spellings in private writing such as letters did not become standardized until a century or so later'. Consequently, the development of an orthography or reforming an existing spelling system can become very contentious because of the highly symbolic relationship between spelling, standard language and community identity. Schieffelin and Doucet (1998: 299) discuss the language ideological debates surrounding orthographic choices for Haitian Creole (*kreyòl*), and Greenberg (2004: 5) describes the efforts in Yugoslav successor states to differentiate their languages from Serbo-Croatian by developing new spelling rules.

Despite the inherent variability in spoken language, standards for pronunciation have also emerged, and the centuries-old tradition of elocution for Standard English (Mugglestone 2003: 14) is evidence of such norms for accent. In the UK, the BBC has been a bastion of spoken Standard English (Schwyter 2016: 3), and the institution only relatively recently embraced regional speech (Crystal 2012: 510). Theodor Siebs (1898) developed the *Deutsche Bühnenaussprache* ('German Stage Pronunciation') for German, which is 'a spoken norm largely based on "educated" and "elitist" opinions of taste' (Mattheier 2003: 230). Even within the same language, there can of course be several standards, as discussed in Section 2.1 with regard to polycentricity. This can apply to spoken standards too, as demonstrated by Austrian standard German (*Hochsprache*), which is used for training newsreaders for the televised national Austrian news (Wächter-Kollpacher 1995: 269–70).

The predominance of the written form for standardization does not extend to all linguistic contexts, such as those where general levels of literacy are low (cf. Abbi 2017; Chapter 4, this volume), where there is widespread heterography (Deumert & Lexander 2013: 528 on languages in Africa) or with regard to pidgins and creoles which normally have no established literary tradition (Siegel 2005: 145–6). As writing practices evolve

along with technological innovations in communication, the categorization of spoken and written becomes increasingly blurred (as, for example, in text messaging) and strictly normative practices apply less and less to the written representation of language (see Section 3.4).

2.2.2 Standard Language and Variability

In 1966, Haugen described codification as 'minimal variation in form' and elaboration as 'maximal variation in function' (Haugen 1966: 931). For Haugen, the ideal case of minimal variation in form would be 'a hypothetical, "pure" variety of a language having only one spelling and one pronunciation for every word, one word for every meaning, and one grammatical framework for all utterances', but the formula of 'minimal variation in form' has nevertheless become part of accepted wisdom about the nature of the standard. Milroy and Milroy (2012: 6) likewise speak of language standardization as involving 'the suppression of optional variability', so non-standard languages permit more variability than standard ones. This has led to discussion as to the extent to which standard languages can embrace variation. Some linguistic aspects of standardization exhibit more variability than others. We have already noted that spelling, being less tolerant of variation and easier to regulate, usually has very little variability, in contrast to spoken norms and the inherent greater fluidity of spoken language. It is also usual to observe greater variability in lexical rather than morphological usage.

Whilst standard languages are often associated with usage in formal contexts, a standard language can comprise various registers, with different levels of formality being associated, for instance, with lexical differences. Trudgill (1999: 120) points out that Standard English encompasses a range of styles, from formal to informal. For example, 'fatigued', 'very tired' and 'bloody knackered' are all Standard English, even if the register of usage varies considerably. In this view, even swear words and slang can be regarded as standard.

Deumert (2004: 3) discusses the division between common conceptualizations of standard/non-standard based on the degree of acceptance of variability and observes that 'the maintenance of variation marks social, ethnic and regional difference', whereas linguistic uniformity is a hallmark of a standard which 'promotes social and political unification and a common identity'. A crucial function of a standard language is to serve as the 'common language' for supra-regional communication, where clarity, consistency and coherence are paramount. This key characteristic is reflected in the names for the standard language in different linguistic contexts, such as *Pǔtōnghuà* (Standard Chinese) and *Kyōtsūgo* (Standard Japanese), which both mean 'common language'. By accepting only one variety as the norm, the *selection* phase of the standardization process strongly implies that deviations from this chosen variety are classified as non-standard. Usage labels in dictionaries such as *non-standard* or *dialect* (both used, for example, in

the Merriam-Webster dictionary) reinforce this view. The categorization of ways of speaking into standard and non-standard can become contentious, as, for example, with the discord identified between 'native' speakers of Breton and new speakers of the more recently developed Neo-Breton standard (Hornsby 2015: 107–9).

In her research on Corsican, Jaffe (1999; also Chapter 16, this volume) builds on Marcellesi's (1989) concept of *polynomie* in her discussion of a polynomic standard for the language, which involves mutual recognition by all speakers and writers of the legitimacy of all varieties of Corsican. This perspective on standard Corsican has helped considerably with tackling feelings of linguistic insecurity that have arisen since the revitalization of the language. Similarly, Frisian is another example of a language with a 'flexible' standard (Feitsma 2002: 213) in the way that it 'leaves room for several different opinions and practices'. It therefore appears to be advantageous for lesser-spoken languages to have standards which are more tolerant of variation and inclusive in their range of norms.

As discussed earlier, pluricentricity also embraces variability by acknowledging multiple standards of what is commonly accepted as the same 'language' (Clyne 1992). On joining the European Union (EU) in 1995, Austria ensured that twenty-three 'specifically Austrian expressions of the German language' were protected in Protocol 10 (De Cillia 1998) as part of the Austrian Standard. Transnational standards of the same language are also recognized (e.g. Trudgill & Hannah 2008; Hickey 2012 for English), and in some cases there is more than one national standard of the same language, as with *Bokmål* and *Nynorsk* in Norway.

2.2.3　Standards and Language Change

Related to the discussion about the amount of variability tolerated by standard languages, there is the question of the degree of *change* standard languages may accept or tolerate. Milroy and Milroy (2012: 19) note that the ideology of the standard inclines us all to view a language as a relatively fixed, invariant and unchanging entity. It is striking that many of the early calls for the standardization of Western vernaculars were motivated by a desire to give them the fixity and stability of classical languages. There is, moreover, plenty of evidence to demonstrate concerted efforts to resist changes to a standard language, as indicated by research on purism (Thomas 1991; Langer & Davies 2005 for Germanic languages; Walsh 2016 for French). Connected with prescriptivism and purism, the 'complaint tradition' (Milroy & Milroy 2012) is also a form of resistance to change because potential innovations are broadly condemned as corruptions to the 'pristine' nature of the standard and would therefore constitute another step towards the language's decline.

Yet, standard languages clearly do change, as Garvin (1959: 28) observed already in 1959, citing 'flexible stability' as one of two intrinsic properties of standard languages (cf. also Davies 2012; Hickey 2012; Curzan 2014). In

practical terms, dictionaries continue to codify new lexemes, as attested by the preface to new editions, such as the Third Edition of the OED, which announces that the number of the dictionary's entries doubled in size from one edition to the next. Spelling reforms are another indicator of the changing nature of standards, although modifications of standard orthography rarely reach completion (if at all) without a great deal of debate and controversy (Sebba 2007: 133–4). Examples since the 1990s are the Czech spelling reform (1994), the German spelling reform (1996–2006), the official switch to the Latin alphabet in Kazakh (2017) and changes to the use of accents in Catalan (2017), along with many other proposed orthographical reforms which never saw the light of day (Sebba 2007: 155).

As we shall see in Section 5, which addresses standard languages in the twenty-first century, the notion of change may refer to shifting conceptualizations of what a standard language should be, rather than simply to changing usage. With destandardization, for instance (Kristiansen 2016), a weakening of the standard language ideology brings with it greater acceptance of variability.

2.2.4   Standardization, Ideologies and Authorities

Since the pioneering study of Milroy and Milroy (2012, 1st edn 1985), a consensus has emerged that it is appropriate to consider standardization as an ideology, and a standard language as an 'idea in the mind rather than as a reality – a set of abstract norms with which actual usage may conform to a greater or lesser extent' (2012: 19). There is a significant body of work on the organization of beliefs on language into ideologies (e.g. Silverstein 1979; Irvine 1989; Blommaert 1999), which may be explicit or tacit (Gal 1989). Nevertheless, discussions of language standards often slip into treating them as a reality. Milroy (2001: 531) himself offers a non-ideological definition of standardization as 'the imposition of uniformity upon a class of objects'. Woolard (1998: 21) notes that, in the vernacular belief systems of much of Western culture, language standards are no longer recognized as 'manmade' constructs, but are naturalized by metaphors such as that of the free market.

Standard language ideologies manifest themselves in the activities of language authorities such as national academies, schools and the media. As Woolard (1998: 7) observes, they are very rarely simply linguistic in nature, but are rather constructed on the basis of social, political and historical factors, and frequently linked to inhabitable positions of social, political and economic power. The chapters in Part II of this volume examine a number of traditional authorities for language standardization – state-appointed institutions such as language academies (Paffey, Chapter 8), metalinguistic texts (McLelland, Chapter 9) and literature (Maraschio & Matarrese, Chapter 11) – but also situations where linguistic authority is uncertain or contested, as in the case of creoles (Migge, Chapter 13) or new speakers (Hornsby &

Ó Murchadha, Chapter 12). In Woolard's terms (2005: 4–5), hegemonic languages rest their authority on a conception of anonymity, which abstracts away from each person's individual characteristics to distil a common, disembodied voice, a 'voice from nowhere'. Conversely, in the case of non-standard, minoritized or regional varieties, the ideology of authenticity locates the value of a language in relation to a particular community, and for the speech variety to be considered authentic, it must be identified as being 'from somewhere' in speakers' consciousness. Standardization of minoritized languages then runs the risk of losing authenticity, which is at the heart of what it is for many to speak a minoritized language. The high ideal of promoting a standard in order to help with the maintenance or revitalization of threatened or endangered varieties may compromise the signalling of identity, intrinsic to authenticity (see Section 3.2.1).

According to Lippi-Green (2012: 173), 'standard language ideology is concerned not so much with the choice of one possible variant, but with the elimination of *socially unacceptable* difference' (original emphasis). In creating stigmatized forms, or supposed 'non-language', standardization can result in 'more (and hierarchical) heterogeneity' (Gal 2006: 21). In Bourdieu's linguistic market (1991: 46–9), the standard language is viewed as a 'normalized product' which 'concurs with the demands of bureaucratic predictability and calculability'. As Coupland and Kristiansen remark (2011: 17), there is thus a tension in sociolinguistic research on standardization. In their words: 'Is linguistic standardisation democratising and in some fundamental sense pro-social, or is it a crude manifestation of social class hegemony?' Moreover, if the standard is becoming more democratic, are the links between standard language processes and exclusion and hegemony becoming attenuated or patchier? In Part V of this volume, questions are raised as to whether in Late Modernity the ideology of the standard is weakening, leading to destandardization, or whether traditional standards are simply being replaced by different, more 'democratic' ones, which may be diffused through online blogs or social media (see Section 5).

3   Beyond Traditional Models and Case Studies

Many of the seminal studies of linguistic standardization have focused on the history of major European languages, underpinned by the monolingual ideology of one language–one nation. As Ayres-Bennett in Chapter 1 demonstrates, traditional models are being reviewed and challenged by opening up the scope and type of case studies. In this section, we review some of the new material and data being analysed and consider the extent to which these are leading to new methodologies and approaches.

3.1   Standardization in Multilingual Contexts

Ulrike Vogl (2012: 37) comments on how forms of multilingual practice are shaped and hidden by the different types of standard language culture. This 'hidden multilingualism' may include multilingualism which is a result of internal variation, language learning or migration. Traditional accounts of the history of major European languages, such as French, have therefore tended towards a narrative stressing the move from diversity to unity, from dialect to standard, from the local to the national. The story is of the spread of French at the expense not just of the dialects, but also of the various regional languages of France such as Basque, Breton, Catalan or Occitan, which inevitably masks the resulting multilingualism. Whilst some speakers will over time inevitably give up their dialect or regional language, others will add the standard to their linguistic repertoire and continue to use one or more other varieties, particularly in more local and domestic spheres.

The monolingual ideology underpinning some of the earlier work on standardization is closely related to the focus on European national languages and the – mistaken – related belief that monolingualism is the norm. Smakman and Nekesa Barasa (2017: 27) observe that the absence of a specific standard language candidate is in fact quite common in non-Western regions due to multilingual situations where more than one language is dominant, either top-down (officially) or bottom-up (culturally). There is thus much to be learnt from looking at other contexts, such as Africa or the Indian subcontinent, where multilingualism has a long history and is not associated with language learning or being part of a social elite. As Nagy and Meyerhoff (2008: 14) contend, 'the world's multilingual speakers should be as much a part of linguistic theory and practice as their more closely-scrutinized monolingual cousins'.

Central to multilingual situations is consideration of contact and code-switching. In Chapter 4, Anvita Abbi explores the contact-induced levelled varieties of Hindi from different regions which are accepted as the standard in spoken forms. This again opens up the question of the relationship between the spoken and the written: in this case, the written and spoken standards differ from each other, yet both are accepted and enjoy prestige in their respective spheres. In Chapter 5, Friederike Lüpke discusses a tension in West Africa between the nationalist views of standardization which seek to protect small languages by giving them the status of written named languages and the fact that this standardization eradicates diversity and undermines the fluid, multilingual use of unmarked code-switching. With this in mind, Lüpke argues for more inclusive models of language development in multilingual settings which embrace rather than efface spoken multilingualism and existing communicative practices.

Finally, a multilingual approach can – and should – in turn be applied to the study of the history of major European languages. Laura Wright (2018), for instance, has challenged traditional accounts of the history of English.