



Universiteit  
Leiden  
The Netherlands

## Transmission and diffusion

Rutten, G.J.; Van de Voorde, I.; Vosters, R.; Condorelli, M; Rutkowska, H.

### Citation

Rutten, G. J., Van de Voorde, I., & Vosters, R. (2023). Transmission and diffusion. In M. Condorelli & H. Rutkowska (Eds.), *The Cambridge handbook of historical orthography* (pp. 596-616). Cambridge: Cambridge University Press. doi:10.1017/9781108766463.030

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3663460>

**Note:** To cite this publication please use the final published version (if applicable).

# Transmission and Diffusion

Gijsbert Rutten, Iris Van de Voorde and Rik Vosters

## 30.1 Introduction

Theories of language change often focus on the various ways in which innovations spread from one language user to another. Following Labov (2007), *transmission* and *diffusion* are two crucial concepts in this context. On the one hand, transmission is used for changes in language passed down and incrementally amplified from one generation to the next, where innovations move in a similar direction across generations. Diffusion, on the other hand, is used to signal changes across communities, with innovations spreading, for instance, from one dialect group or one part of a language area to another, and involving imperfect adult language learning. Since their introduction, both concepts have gained wide currency in sociolinguistics, but the implications and relevance of this twofold distinction for orthographic change have not been assessed in great detail.

Exploring the relevance of the transmission versus diffusion dichotomy is exactly what we set out to do in this *Handbook* chapter. Departing from the classic definitions of transmission and diffusion by Labov in Section 30.2, we evaluate their applicability to diachronic orthography, comparing them to other, related concepts, to a large extent based on the work of Milroy (2007). We then move on to discuss examples of transmission (Section 30.3) and diffusion (Section 30.4). In Section 30.5, we shift our attention to another layer of orthographic variation and change, above the notion of the community that underpins transmission and diffusion, namely the layer of multiple standards in situations of pluricentricity. We then move to a broader reflection comparing diachronic perspectives on spelling variation and change to other instances of language change, while discussing the issue of transmission versus diffusion within the larger framework of standardization and the development of linguistic norms

within and across speech communities over time (Section 30.6). Section 30.7 concludes.

In line with our own research expertise, we adopt a historical sociolinguistic approach to the topic outlined above, delving into several examples of regionally or socially motivated language change. As such, we do not only look at the peculiarities of spelling as a learned behavior, but also explore the link between orthographic change, diffusion and supralocalization, which offers clear implications for the study of standardization as a historical process. We argue against a ‘tunnel view’ of the standardization process, where regional writing traditions lead to one clear and exclusive national standard. Instead, we focus on how such regional standards actually give rise to a more multilayered sociolinguistic landscape, as is the case for many pluricentric languages. To evaluate the different theoretical concepts discussed in this chapter, we draw on different case studies from our own research on Dutch language history, but also from case studies in German and English historical sociolinguistics. By taking this cross-linguistic perspective, we explore to which extent issues related to the diffusion and transmission of orthographic change are similar across languages, pointing to interesting avenues for comparative historical sociolinguistic work and applications to other languages and language families.

## 30.2 Transmission, Diffusion and Writing Traditions

The two main models of language change discussed by Labov (2007) are transmission and diffusion, which differ in a number of ways. The most important difference is the type of language learning implied in the model. Transmission involves native-language acquisition by children, ensuring continuity across generations. Children are seen as capable of replicating faithfully the language of their parents’ generation so that the phonological system – Labov’s theory primarily deals with sound change – of each new generation copies that of the previous generation. This does not exclude change, quite the contrary in fact, but where change occurs it is usually in the same direction across the generations, for example, by generalizing an innovation to more phonetic contexts than before. Labov (2007: 246) calls this process *incrementation*. Transmission and incrementation thus lead to internal change, that is to say, change from within the system of the language, also termed change ‘from below’: properties of the system are transmitted to new generations who may increment these to new contexts. This first model of language change ties in with the tree model used in historical linguistics (Schleicher 1853), which also works on the combined assumptions of intergenerational transmission and increasing difference to other languages as a result of internal change. It may be clear that this model is primarily and perhaps only applicable to relatively stable and homogenous speech communities.

Social changes and occurrences such as famine, war and migration may render communities less stable and less homogenous, and may foster contact between speech communities. In such contact situations, language change occurs through the diffusion of forms from one community to the other. Since this involves the importation of forms from other language systems, language change occurs ‘from above’. In the prototypical case, this involves adult learning, which is often imperfect. This means that specific sounds, for example, may be borrowed, but the intricate phonological system ruling these sounds in the donor language will remain obscure to the adult learners, as such intricacies are generally only learnable by children. In other words, only some properties diffuse. This kind of diffusion across communities, as opposed to transmission within communities, ties in with the wave model in traditional historical linguistics (Schmidt 1872).

Labov’s approach has given rise to much research on innovation and change, largely on synchronic data, and it has subsequently been criticized and refined, for example by Cheshire et al. (2011), Stanford and Kenny (2013), Tagliamonte and Denis (2014). Cheshire et al. (2011: 156) consider it to be a useful toolkit for the study of language change, and this is also the perspective that we take. We ask whether transmission and diffusion, though primarily based on sound change in the present, are useful concepts in the study of diachronic orthography. One could argue that this is a curious question, as orthography – and more generally writing – is a much more conscious act than speaking, and moreover not part of natural language acquisition but exclusively learnt through instruction (Coulmas 2013: 98). From this observation, it would follow that the distinction between internal changes ‘from below’ and changes ‘from above’ does not hold: all orthography being ‘from above’, diffusion would be the only concept needed. The orthographic systems of European languages are furthermore not bound specifically to their respective communities. On the contrary, many communities have adopted the Latin script and Latin-based phoneme–grapheme correspondences, after which the writing system was adapted in order to represent sounds not part of the Latin phoneme inventory (Coulmas 2002). In this sense, orthography can be seen as one colossal contact phenomenon, as a technique imported through diffusion from another language. We argue, however, that orthographic change can often be modeled in much the same way as ‘normal’ language change: it frequently follows an S-curve, for example, and is subject to a range of internal and external factors (see Sections 30.4 and 30.5). As we discuss, orthographic practices are also not only the result of explicit instruction, but can equally be picked up during reading in a less conscious manner (see Subsection 30.3.1). It does therefore not seem justified to treat orthography differently from other language levels *a priori*. In addition, we argue in Section 30.3 that instances of transmission can also be found in the history of orthography.

Even if we treat orthography as a regular language phenomenon akin to phonology and morphosyntax, Labov's approach comes with one serious drawback. The distinction between transmission and diffusion is based on the ability to define concrete and separable language communities, that is communities of language users with shared sets of norms. Transmission and incrementation take place within such communities, whereas diffusion is a contact phenomenon transcending the limits of communities. The two processes can be distinguished particularly in comparative analyses of well-defined and fairly homogenous communities with different migration and settlement histories (Tagliamonte and Denis 2014). But communities themselves can be multilingual, bringing in language contact as an intrinsic property (see Cheshire et al. 2011). More generally, contact should be treated as a normal phenomenon in language histories affecting any community, if only because early and late modern population growth in northwestern Europe can often only be explained by in-migration, as until recently death rates outnumbered birth rates, particularly in urban areas (Milroy 2007: 150). At the same time, delineating language communities can be challenging in the case of historical periods for which sufficient contextual information is lacking. Labov (2007: 348) also acknowledges that his concept of the community may be strongly influenced by the present-day North American communities that he has worked on, and he points out that European dialectology usually focuses on the diffusion of variants in historical dialect areas, from one region to the other, less so on transmission through incrementation.

Milroy (2007) introduces a distinction similar to Labov's (2007) transmission and diffusion, while placing contact at the heart of the matter. Diffusion can be compared to *supralocal* changes, which are available 'off the shelf': they seem to travel easily from community to community, over large territories, and are picked up by speakers across the language area, or even across the globe (see also Cheshire et al. 2011). This also means that direct contact between speakers is not necessary. The sociolinguistic profile of such changes is still under-researched (Milroy 2007: 162), but obvious present-day examples are popular expressions such as *be like* (see Tagliamonte and Denis 2014). Transmission occurs within *local*, close-knit networks and communities, where intense, direct and ongoing contact enables speakers to replicate the fine-grained constraints that regulate the variable system. Milroy and Milroy (1985b: 363) have emphasized the importance of language *maintenance* in such local, high-contact settings. Local changes occur when the network structure changes, or through contact with other networks and communities via peripheral network members with relatively loose ties.

Labov (2007) focuses on the difference between children's and adults' ability to learn linguistic rules. Milroy's (2007) view also has implications for acquisition, but it is not a central theme, since maintenance within networks affects both children and adults. The Milrovia contact-centered view

does not assign different models of language change to different age groups. Interestingly, Stanford and Kenny (2013) argue that density of interaction is in fact sufficient for a computational model of Labov's (2007) results, and can explain the outcomes of transmission and diffusion without taking into account possible language acquisition differences between children and adults. Another reason to not accept the difference between children and adults lies in the inherent reversibility of orthography: anyone can take the time to learn the phonetic and/or phonological intricacies underlying a specific spelling system, and usually language users can return to their writings in order to amend the spelling. There is no critical threshold for the acquisition of orthographic rules, as other factors can, of course, also play a role in the acquisition of orthography.

For the present chapter, we depart from the notion of *writing traditions*, particularly from regionally bound orthographic practices emerging in the course of the medieval period and undergoing changes throughout medieval and postmedieval times. In European settings, among the most pervasive and significant changes are without any doubt *supralocalization* and subsequent *standardization* (Nevalainen and Tieken-Boon van Ostade 2006, Rutten 2016b), understood as the increase of the geographical span of certain variants at the expense of other variants, moving toward a more focused variety (Le Page and Tabouret-Keller 1985).<sup>1</sup> In line with the European research tradition, we take *diffusion* to be the main source of orthographic change, defining it in general terms as the spread of innovations through the language area as a result of contact (and not restricting it to adult learning effects across communities). We avoid *transmission* in the traditional sense as much as possible as it depends on the debatable notion of homogenous communities, although we also claim that it may be a useful concept in specific sociohistorical contexts: these include explicit instruction in schools, for example, in which orthographic norms are consciously transmitted to the next generation, but also in close-knit networks of adults. We prefer to call this phenomenon 'transmission' instead of 'maintenance', as it often involves educational settings, and also because new orthographic norms may be introduced and transmitted before any kind of maintenance is logically possible.

### 30.3 Transmission

#### 30.3.1 Transmission through Instruction

As we already discussed, in general, orthography represents explicitly learnt behavior. Although there are language users who largely acquire reading and

<sup>1</sup> See also Chapter 13, this volume, for a discussion on standardization and supralocalization. The latter is also discussed in Chapter 21, Chapter 25 and, although briefly, Chapter 4, this volume.

writing skills independently through self-teaching, the acquisition of reading and writing usually involves explicit instruction by someone who is already experienced. The key issue here is the transmission of conventions, without which continuity in the written tradition would not exist. Like today, transmission of orthographic conventions was institutionalized in medieval and postmedieval European history, with monasteries and schools providing the typical contexts where children, adolescents and adults learnt to spell (Boekholt and De Booy 1987: 1–9). In addition, home teaching resulted in reading and writing skills among a larger part of the population, including women, although it should be noted that many places also offered schools for girls (Boekholt and De Booy 1987: 30).

Apart from the observable continuity in writing traditions, another piece of indirect evidence of transmission comes from the extensive metalinguistic traditions found all over Europe. The Dutch, German, English, French, Italian, Spanish and Portuguese grammatical traditions begin in the sixteenth century, or just a bit earlier, and orthography is a prominent topic therein (Carvalho Buescu 2000 for Portuguese, Dobbins 2000 for Dutch, Esparza Torres 2000 for Spanish, Kristol 2000 for French, Marazzini 2000 for Italian, Rössing-Hagar 2000 for German and Verrac 2000 for English). The sources testify to the practice of orthographic transmission, but the general problem with analyzing transmission in this way is similar to the problem connected to the institutional and family settings mentioned above: we can assume metalinguistic texts played a role in the transmission of orthographic conventions, as we can assume children and adults learnt to read and write through instruction by parents and peers, but in general we lack detailed information about the exact ways in which conventions were transmitted to new language users. For example, we do often not know whether extant metalinguistic texts were actually used in the historical classroom, and if so, which texts exactly, and to what extent and how the teacher made use of them (Vandenbussche 2007b: 29). With respect to the normative grammatical tradition of the early and late modern period (approximately sixteenth to nineteenth centuries), research is in fact quite skeptical of direct influence on patterns of language use (see Rutten et al. 2014, and Anderwald 2016 and the references therein).

The observable continuity in orthographic conventions in itself provides evidence of transmission, but the conclusion, for example, that <f> and <v> have been used for labiodental fricatives throughout history, does not generate much insight into how this convention was transmitted. Nonetheless, discontinuity or divergence from common norms can indirectly reveal teaching practices. The variable use of <h> in Middle English sources, including both deletion of expected <h> and prothesis of <h> in vocalic onsets, strongly suggests that speakers who lacked the phoneme /h/ had been taught to write <h> or had seen it written in words where they would normally not pronounce /h/. Milroy (1992a: 199) lists many thirteenth-century examples from East



Anglia including forms of the verb ‘to have’ such as *adde*, *adden*, *aue*, and also *ate* for *hate*, *is* for *his*, *om* for *home* and many more, as well as the addition of ‘unhistorical’ *h* in words such as *ham* for *am*, *his* for *is* and *hure* for *our*. This kind of discontinuity, resulting from a conflict between pronunciation and orthographic conventions, can itself be quite continuous: Hundt (2015: 97), for example, discusses words such as *as* for *has*, *ealth* for *health*, and also *ham* for *am* and *hare* for *are* in English private letters from 1846.

Similar instances of *h*-dropping and *h*-prothesis can be found in Middle Dutch texts, in particular from the southwest of the language area such as Flanders, and these writing practices also appear to have been continued into the early modern period. Rutten and van der Wal (2014: 24–25) report on a range of variable features found in seventeenth-century private letters from the Zeeland region in the southwest of the present-day Netherlands. Apart from prevocalic deletion of <h> as in *andt* for *handt* meaning ‘hand’ and prevocalic prothesis as in *hacht* for *acht* meaning ‘eight’, they mention two additional and related effects of teaching an orthographic system with <h> to language users who lack /h/ in the spoken language. One of the historical Dutch pronunciations of the letter <h> is [ha:], rendered [a:] by speakers lacking /h/. The letter <a> is, however, also rendered [a:]. This means that for *h*-less language users, the alphabet provides two graphemes with the same phonetic value, which led seventeenth-century letter-writers from Zeeland to substitute <a> for <h> as in *aebben* instead of *hebben* ‘have’, as well as use <h> for <a> as in *hpril* for *april* ‘April’.

These English and Dutch examples demonstrate divergences from what was presumably transmitted, and they thus provide indirect evidence of teaching practices. There are, however, also historical sociolinguistic situations where explicit transmission can be better reconstructed, together with its effects on patterns of language use. What is needed is (1) evidence of actual teaching practices, and (2) a well-defined, real or imagined language community. To measure the effect of explicit transmission, we also need (3) a balanced corpus of original sources. Requirement 1 is without doubt the hardest to meet. Archival sources fulfilling requirements 1 and 2 are discussed, for example, by Van der Feest Viðarsson (2017), who analyzes Icelandic student essays from the nineteenth century. These are produced within a well-defined educational community setting, and corrected by the teachers, providing evidence of corrective attitudes and practices.

An excellent opportunity to study transmission through instruction is provided by the new Dutch language policy developed around 1800 in the context of nation-state building, which resulted in official rules for the spelling and grammar of Dutch, published in 1804 and 1805 (Rutten 2019). At a time when the nationalist ideology started to supersede older, regional identities and was politicized into a nation-building project in the first decades of the nineteenth century, the new language policy was connected to the



simultaneous reorganization of the field of education, which led to a national system of school inspection (Schoemaker and Rutten 2019). The extensive archives of the school inspection show a profound awareness of the language policy among teachers and school inspectors alike, and provide evidence that, already in the first decade of the nineteenth century, many schoolteachers made an effort to transmit the newly devised prescriptions to their pupils. Schoemaker and Rutten (2019) and Rutten et al. (2020) show that schoolbooks almost immediately shifted to the new orthographic norms, again suggesting a high awareness of these new norms. The explosion of new schoolbooks and of revised editions of old books also implies general exposure to the new orthography. Furthermore, corpus studies of private letters, diaries and regional newspapers from before and after the new policy show a great influence of the new orthography on the writing practices of individual language users, even in private writings (Krogull 2018). We can assume language users adhered to new orthographic norms both to show erudition and to signal loyalty to the nation-building endeavor. For a range of spelling variables, an almost complete shift from older conventions to the new prescriptions can be found. The fact that newspapers quickly shift again implies that exposure to written language often implied exposure to the new rules (Krogull 2018).

The Dutch case of the decades around 1800 shows that the transmission of orthographic norms can be fruitfully studied if archival sources are available providing evidence of norm transmission as well as norm acceptance in language use (requirements 1 and 3 mentioned above). The ideological background is the official effort to create, through the educational system, a homogenous community constituted by the inhabitants of the Dutch nation-state (requirement 2).

### 30.3.2 Transmission within Networks and Communities

As explained in the previous subsection, we do not limit transmission to child language acquisition. We should rather investigate whether certain adult contact situations outside the educational settings discussed in Subsection 30.3.1 may count as instances of transmission. Close-knit networks and communities constitute possible sociohistorical contexts in which orthographic practices are transmitted from one language user to the other, either as instances of orthographic maintenance or as examples of orthographic change (requirement 2). In the absence of explicit instruction, requirement 1 cannot be met, which shifts the burden of proof to empirical sources demonstrating norm enforcement (requirement 3). Nevalainen and Raumolin-Brunberg (2016: 220–21) discuss language variation and change as found in English letters related to the sixteenth-century Johnson merchant family network. They show how the language use of some individuals within this network seems to pattern similarly with respect to six morphosyntactic

changes in progress (e.g. use of single negation); they do not discuss spelling. Interestingly, the similarities appear to follow the social pattern of master-apprentice bonds. The sociolinguistic profile of the London-based merchant, John Johnson, is closely matched by his apprentice based in Calais, Henry Southwick. Likewise, Johnson's relative Anthony Cave has a profile similar to that of his apprentices Richard Johnson and Ambrose Saunders. Nevalainen and Raumolin-Brunberg (2016: 221) argue that this shows that generational transmission not only takes places from parents to children, but also from masters to apprentices.

Studies that focus on spelling transmission include Sairio (2009) and Conde-Silvestre (2019). Sairio (2009: 226–92) seeks evidence of network-based orthographic similarities in the eighteenth-century Bluestocking network around Elizabeth Montagu. An important conclusion is that correlations between network strength and spelling choices are not strong, and that other factors such as social position, gender and the type of relationship between writer and addressee are more important. Conde-Silvestre (2019) analyzes the fifteenth-century Stonor letters, a collection of English letters centered on Sir Thomas Stonor II. Conde-Silvestre (2019) argues that the letter-writers constituted a community of practice (CoP) (Eckert 2000) as they were all engaged in the same endeavor of maintaining the feudal system known as *enfeoffment*. He compares the spelling in these letters to a related collection of letters, also part of the Stonor collection, though not produced by people participating in the same system of *enfeoffment*. For a number of spelling variables, Conde-Silvestre (2019) shows that the language of the CoP is more focused, and thus less variable, clearly suggesting orthographic norm enforcement.

Metalinguistic networks and communities constitute another context in which orthographic convergence may occur. It is traditionally claimed that medieval scribes and early modern printers advanced spelling convergence (see also Voeste 2007a). Rogos (2013), Rutkowska (2013b) and Tyrkkö (2013) demonstrate that this convergence can be interpreted as a function of the scribes and printers being part of communities of practice. Tyrkkö (2013) analyzes 124 English books published by 88 different printers between 1500 and 1700, and shows a remarkable decline of spelling variation. Rogos (2013) focuses on fifteenth-century scribes copying one of Chaucer's *Canterbury Tales* and the "centrifugal tendencies" (Rogos 2013: 118) characterizing consonant spelling and abbreviations. Rutkowska (2013b) makes similar claims for sixteenth-century printers based in London. She also argues that in order to account for focusing, also diachronically, we need to consider that early printers operated within communities of practice at three different levels: at the most local level, people collaborated within printing houses; in addition, different printers were active at the same time, part of the same guild, and often had their offices in the same street or neighborhood; finally, within

these communities, particular ways of presenting texts were also transmitted to younger generations.

A fruitful way to study metalinguistic transmission is by focusing on variants that were nonexistent or very low-frequency at the time when they were prescribed in metalinguistic publications. McLelland (2014: 266) shows that the German seventeenth-century grammarian Schottelius advocated *kk*-spellings in words such as *wekken* ‘to wake’ (compare present-day German *wecken*). For about one generation, quite a few elitist language commentators and literary authors followed this practice, but by the final decades of the seventeenth century, *kk* had already disappeared again from spellings like the example above. Flemish grammarians of the eighteenth century represent a similar case (Vosters et al. 2012, Vosters and Rutten 2015). They prescribed a complex system of diacritics to distinguish etymologically different *e*- and *o*-sounds. Here, too, the spelling was taken over primarily by other language commentators and a limited number of printers. Such cases are indicative of transmission processes within professional communities.

## 30.4 Diffusion

So far, we have singled out certain sociohistorical contexts that may be analyzed from the perspective of transmission. In this section, we focus on the more general process of diffusion, in particular on three main types, that is diffusion across regions, across social groups, and across genres and contexts. We limit ourselves, in the context of this chapter, to a few illustrative examples. Since language is seen as a variable social system in our historical sociolinguistic approach, a main concern for diachronic orthography is to gain access to sizable corpora of historical sources, as equally as possible distributed across social variables such as region, social rank and gender. In addition, it is important to compare similar genres and/or registers, produced in similar contexts and with similar purposes, as it is well-known that different spelling systems have historically been used for different genres and/or registers: Salmon (1999) discusses differences between print and manuscript spelling in the history of English, and Rutkowska and Rössler (2012: 225–26) refer to the so-called ‘double spelling standard’ in eighteenth-century English, when authors, grammarians included, used different orthographies in their private letters and printed texts.

### 30.4.1 Regional Diffusion

Studying regional diffusion requires substantial sources from various regions, preferably from similar genres, and where possible from different times. Relative proportions of variants at one moment in time can of course, in

line with the apparent time construct, indicate an ongoing change, but one advantage of historical sociolinguistics is that it is possible to study language change in real time. Focusing on printed sources, Rössler (2005) analyzes regional diffusion in a corpus of texts originating from Bavaria, in southern Germany, and Austria. The corpus comprises various genres, organized into two major types: religious texts and secular texts. The sources date back to the sixteenth, seventeenth and eighteenth centuries, when religious conflicts and, in connection with these, religious texts were significant, and when orthographic supralocalization occurred along with metalinguistic reflection, while regional writing traditions remained important. The sources are localizable based on the places where they were printed; the cities taken into account are Vienna, Graz, Klagenfurt, Salzburg, Linz, Innsbruck, Munich and Passau. Rössler explains that many more historical actors were involved in text production than just the author: books were usually printed in offices with multiple typesetters and proofreaders so that the number of people responsible for the language of the final product exceeded the number of authors. Rössler even argues that typesetters and proofreaders influenced the spelling and morphology of a text more strongly than the author, at least until the individual author gained significance as a cultural figure from the middle of the eighteenth century onward. The text samples in the corpus are 4,000 to 5,000 words on average, per genre, per three-decade period and per city, resulting in a corpus of 437,279 words.

The principled setup of the corpus allows Rössler to perform detailed analyses of the gradual reduction of variation characteristic of the period. This reduction takes place at various levels: the number of competing variants decreases, but also the number of competing regional writing traditions. The latter development, in particular, signals supralocalization. Crucial in the analysis of such diachronically diverse writing traditions is the empirical baseline evidence against which the orthographic changes are measured. This can only be a reliable historical phonological reference system (see Elmentaler 2003: 97). On the basis of such a system, Rössler helpfully distinguishes between various levels of analysis. The lowest level is that of the concrete *variants* found in textual sources. Texts are often not homogeneous, displaying several variants for one phoneme. The individual variants of a text taken together constitute its *writing variety* (*Schreibvarietät*). The aggregate of multiple localizable writing varieties taken together constitute the *writing variety of an area* (*Schreibvarietät eines Raums*), which we can also call *regional writing traditions*. Regional diffusion in this case means that a variant or the variants used in a given text are considered to be part of a different regional writing tradition than the text itself.

While regional diffusion in orthography may be seen as a phenomenon characteristic of medieval and early modern times, pre-dating the advent of standardization impulses, various studies of historical Dutch have shown that

particularly manuscript sources from more recent times may also show intriguing patterns of regional diffusion, often in direct connection with standardization and/or language planning. Vosters et al. (2012) and Vosters and Rutten (2015) show how Flemish administrative sources from the early nineteenth century display an orthographic shift within only six years, namely between 1823 and 1829. These years represent the beginning and the end of the Dutchification policy during the United Kingdom of the Netherlands (1814–30). The oldest sources have an average of about 42 percent of spelling variants usually considered typically southern, that is, Flemish. By 1829, after just six years, these variants have decreased to a mere 24 percent in usage, while the typically northern, that is, Dutch forms have diffused further among southern writers.

Regional diffusion often implies the introduction of graphemes and/or grapheme–phoneme correspondences exogenous to the region to which the variants are diffusing. The result is a further widening of the gap between the spoken and the written language. Rutten and van der Wal (2014: 19–74) demonstrate this on the basis of an analysis of the interplay of phonology and orthography in the *Letters as Loot Corpus*, comprising seventeenth- and eighteenth-century Dutch private letters. One example concerns the representation of etymologically different *e*-sounds, where one area, namely Zeeland, maintains the historical difference between so-called sharplong *e*, deriving from Germanic diphthongs, and the so-called softlong *e*, which are the product of vowel lengthening, while another region, that is, Holland, has a merger. Various regional writing traditions coexisted in the seventeenth and eighteenth centuries: a southern tradition incorporating the etymological difference in the spelling, and various northern traditions indicating the merger by adopting morphemic or syllabic spelling principles (see Rutten and Vosters 2013). The southern writing tradition was backed up by metalinguistic discourse and became officially standardized in 1804 (see Subsection 30.3.1). In the meantime, however, many letter-writers from Zeeland had shifted to the merging spelling systems characteristic of Holland. The example shows that the regional diffusion of orthographic practices can lead writers to disregard phonological differences that they maintain in their spoken language even if presumably prestigious metalanguage promoted an orthography in line with their speech.

#### 30.4.2 Social Diffusion

Older manuscripts, such as medieval administrative texts, are often localizable and thus enable research on regional diffusion. Social diffusion is a more difficult topic when it comes to the availability of representative sources. It is not only essential to identify the actual or main author of a text, in the case of multiple authorship; it is also necessary to identify the author in terms of

relevant social characteristics such as position or rank within a network, community or in society at large, level of education, degree of participation in the written culture, gender, age or generation, religious beliefs and so on. This requires ardent sociohistorical research and, admittedly, some sheer luck. A diversity of social characteristics in the sources, imperative for a comparison across social variables, can be reached by focusing on texts ‘from below’, that is on genres that were widely produced and used by large portions of historical populations, also from more humble social backgrounds, such as letters and diaries (Elspaß et al. 2007, van der Wal and Rutten 2013). More easily available sources, such as literary texts, are often bound to specific social groups, for example well-educated men. However, texts ‘from below’ were more often reused or thrown away, and their authors have left fewer traces in the archives, which impedes research on their social characteristics (Elspaß 2012b: 159).

Rössler (2005), discussed above, circumvents the issue of authorship by focusing on the place of printing. This enables him to analyze the influence of religion on orthographic choices. A major change in his materials is the gradual disappearance of Upper German spellings in his corpus of Upper German texts, signaling the ongoing supralocalization in German orthography. This change is, however, halted or temporarily reversed in the seventeenth century during the intense Catholic Counter-Reformation. In this specific socio-historical context, a regional writing tradition with Upper German elements came to index locally supported Catholicism. Later in the century, this effect disappeared, and supralocalization advanced.

When there is enough information about authors and their social profile, meaningful analyses of the interplay of social variables and orthographic practices can be undertaken at the level of individual language users. Elspaß (2005) presents research on a corpus of 648 German private letters amounting to about 375,000 words. Most letters are so-called emigrant letters sent from and to North America in the nineteenth century following the migration waves. Elspaß could determine the educational background of quite a few writers in the corpus, adopting a two-way distinction between those with solely primary and those with primary and secondary education. For a number of orthographic variables, Elspaß is able to show that this social difference marks a considerable divide in spelling. Spelling variants such as *wier*, *mier* and *dier* for *wir*, *mir*, *dir*, for instance, almost never occur in the writings of scribes with a higher than elementary education, while they do occur to varying degrees (between 7 and 17 percent of all variants) in the writings of less educated scribes (Elspaß 2005: 439).

An even more fine-grained social division is part of the aforementioned *Letters as Loot Corpus* (Rutten and van der Wal 2014). This corpus is socially stratified in terms of gender and social rank. In addition to considerable numbers of letters written by women, it also comprises letters associated with four

different social layers in contemporary Dutch society. Rutten and van der Wal distinguish between the lower, lower-middle, upper-middle and upper ranks, largely based on education and occupation, and ranging from hardly educated sailors and their wives in the lower ranks to established captains and their wives in the upper ranks. In this corpus, punctuation is a relatively marginal phenomenon in two respects. More letters lack than contain punctuation, and if punctuation is used, it is often inconsistently and not throughout the entire letter (Rutten and van der Wal 2014: 268). On the basis of the social stratification of the corpus, it can be shown that punctuation in private manuscript sources rises in this period as a social change ‘from above’. Only 27 percent of the seventeenth-century letters have some kind of punctuation, which increases to 42 percent in the eighteenth century. Crucially, upper and upper-middle ranked men are the leaders of this change with c. 60–70 percent of letters with punctuation already in the seventeenth century. They are closely followed by upper-ranked women, whose use increases from 14 percent in the seventeenth to over 60 percent in the eighteenth century. Among the lower and lower-middle ranks, punctuation is still infrequent in the eighteenth century. Rutten and van der Wal (2014: 269) interpret these social constraints of rank and gender as symptoms of the same social variable, namely writing experience, taken as the degree to which individuals participate in the written culture.

### 30.4.3 Diffusion across Genre and Context

When studying regional and social diffusion based on the regional and social characteristics of text producers, such as printers and writers, it is important to keep in mind that genre can be a crucially intervening factor. The genre will normally not vary in the case of diachronic comparisons, while synchronically, comparisons across genres are indeed quite informative. Classifications of genres, text types and also situational contexts will often depend on the specific sociohistorical situation from which the sources emerge. Stenroos (2004) presents a detailed analysis of the spread of the spelling <th> in Middle English based on fourteenth- and fifteenth-century data taken from *A Linguistic Atlas of Late Mediaeval English* (LALME, McIntosh et al. 1986a). The spelling <th> was and still is used for dental fricatives, both voiced and unvoiced, in initial, medial and final position (e.g. *think*, *either*, *both*). In the period under investigation, it competed with the older letter thorn, <þ>, and the more recent innovation <y>. Stenroos (2004) distinguishes various systems used in northern and southern areas, some with graphemic mergers, others which maintain differences between the graphemes. One factor influencing the choice of grapheme is voice: <th> occurs first in voiceless positions, only later in voiced ones. The overall picture, thus, is one of regional diversity and graphemic as well as phonetic conditions. But Stenroos (2004)



also shows that genre is a major factor. Grouping the various texts in LALME into two major genres, that is, documents and literary texts, it turns out that southern documents are very progressive with almost 100 percent instances of <th> by the middle of the fifteenth century, when the use thereof is still fairly marginal in literary texts. In northern areas, <th> is well advanced in voiceless positions, both in documents and literary texts. It is a minority form, however, in voiced contexts. Stenroos then argues that the spread in the north and the south, though different in terms of pace, and of graphemic and phonetic conditioning, was crucially associated with genre. In particular, the use of <th> diffused from legal texts from the London/Westminster area to legal texts elsewhere as a function of the centralized legal education system in Westminster.

Genre-related variation persists in the early and late modern period, even in administrative sources at a time when standardization gains high significance. The aforementioned corpus of Flemish administrative sources from the early nineteenth century comprises two variational dimensions related to genre and context (Vosters 2011, Vosters et al. 2012). The first relates to genre or text type in a conventional way: the three main types in the corpus are police reports, interrogation reports and indictments, where indictments clearly constitute the most formal type. The proportion of typically southern/Flemish spellings is c. 30 percent in the police and interrogation reports, but drops below 20 percent in the indictments. The second dimension concerns the institutional level at which the texts are produced, and so is more related to context than to genre *per se*. The results are similar in that locally and regionally produced sources display c. 30 percent typically southern/Flemish spellings, which drops below 20 percent in supraregional sources (Vosters 2011: 319–20).

## 30.5 Diffusion and Pluricentricity

### 30.5.1 Introduction

Above (Section 30.4), we have discussed examples of diffusion at different levels and different stages: we evidenced how emerging regional writing traditions spread geographically in the process which we labeled supralocalization, which then provided impulses for processes of standardization at an even larger level. However, adopting a more macro-level perspective, we can also observe how processes of supralocalization and ensuing standardization, in combination with the emergence of regional and (proto)national borders, may lead to a new situation: multiple centers may emerge from which norms are diffused, thus giving rise to distinct yet related national or regional language norms. Such is the case for pluricentric languages, which are “used in more than one country as national or regional official language” and where

this situation “has resulted in differences in the standard variety” (Ammon et al. 2016: xxxix). We believe that taking a pluricentric perspective on language history can help to combat the so-called “tunnel vision version” of language history (Watts and Trudgill 2002: 1; see also Watts 2011: 585): a highly teleological view presenting “language history as the inexorable march towards a uniform standard” (Elspaß 2007: 4). Such a teleological view generally implies a focus on regional variation in the medieval period, which gives way to an exclusive focus on standardization and a unified standard language as the clear and exclusive outcome of this process from early modern times onward. Historical sociolinguistic research (see Rutten 2016b) has advocated a shift away from this tunnel view of standardization, in favor of a more nuanced and multilayered perspective, in which we give prominence to ongoing variability in written language use. In fact, in many cases, we argue, one can observe how regional writing traditions, instead of leading to a sole and exclusive national standard, in fact evolve into a pluricentric landscape. This is the case for the standardization history of Dutch, which we discuss in more detail as our primary case study by focusing on diffusion and historical pluricentricity across the centuries.

### 30.5.2 A Case Study: The Diffusion of Long *a*

Our case study concerns the orthographic representation of long *a* in closed syllables. For this variable, the traditional practice, originating from Middle Dutch (approximately twelfth to sixteenth centuries) spelling praxis, was to indicate lengthening by adding an <e> (rarely <i>) to the original vowel <a>, resulting in spellings such as *maend* ‘month’ and *daer* ‘there’. From the seventeenth century onward, however, this older writing practice was increasingly being replaced by the doubling of the original vowel as the incoming form, largely irrespective of the pronunciation of the long *a*. This resulted in <aa> rather than <ae> spelling, for example *maand* ‘month’ or *daar* ‘there’.<sup>2</sup> As such, we focus on how the incoming orthographic variant <aa> was diffused across the language area instead of older <ae>.

To investigate this orthographic change, we used a preliminary test version of the multigenre *Historical Corpus of Dutch*, which was developed at Vrije Universiteit Brussel and Leiden University. This corpus includes texts from the middle of the sixteenth, seventeenth, eighteenth and nineteenth centuries, and the preliminary version we used comprises a component of pamphlets and a component of so-called ego-documents. Pamphlets are printed texts, mostly commentaries or polemics about current affairs, political or religious topics, as well as some public ordinances and similar documents.

<sup>2</sup> Some other spelling variants also existed (e.g. <a> spellings), but they hardly played any role from the early modern period onward. For this reason, we limit our attention to the two main variants, <ae> and <aa>.

Ego-documents, on the other hand, were not meant to be published: they are more personal, handwritten texts, such as diaries, travelogues and personal chronicles of local events or family history, often primarily produced for the authors and their family members (see Elspaß 2012b). The corpus was built with the aim of studying pluricentricity in Dutch language history, and as such it includes material from both the central area (Holland) and a more peripheral region (Zeeland) in the north of the language area (i.e. the later Kingdom of the Netherlands), as well as material from the center (Brabant) and the periphery (Vlaanderen) in the south of the language area (i.e. the Dutch-speaking territories of what would be later called the Kingdom of Belgium). While the northern province of Holland became the center of standardization from the seventeenth century onward, Brabant remained a central and influential region in the south.

The diachrony of the ongoing change from <ae> to <aa>, as shown in Figure 30.1, tells us that the incoming form first appears in the mid-seventeenth century, gains significant ground by the mid-eighteenth century, and clearly becomes dominant by the mid-nineteenth century. Overall, of all tokens ( $n=4,745$ ), the incoming variant accounts for 35.3 percent in the ego-document subcorpus ( $n=827$ ), whereas it accounts for 37.4 percent in the pamphlet subcorpus ( $n=898$ ): as such, the pamphlets are slightly quicker in adopting the new form, but this difference is really small. More generally, however, the incremental change from <ae> to <aa>, which visually appears

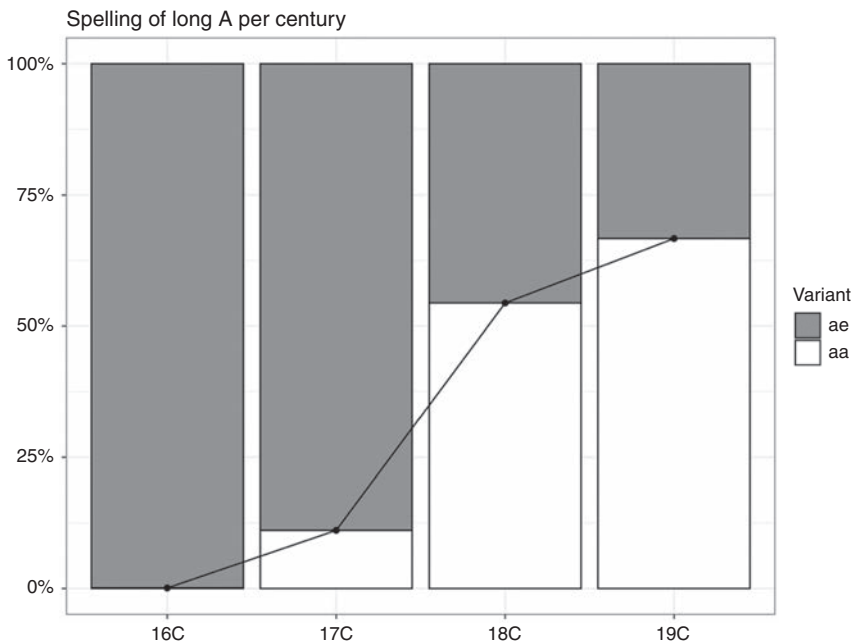
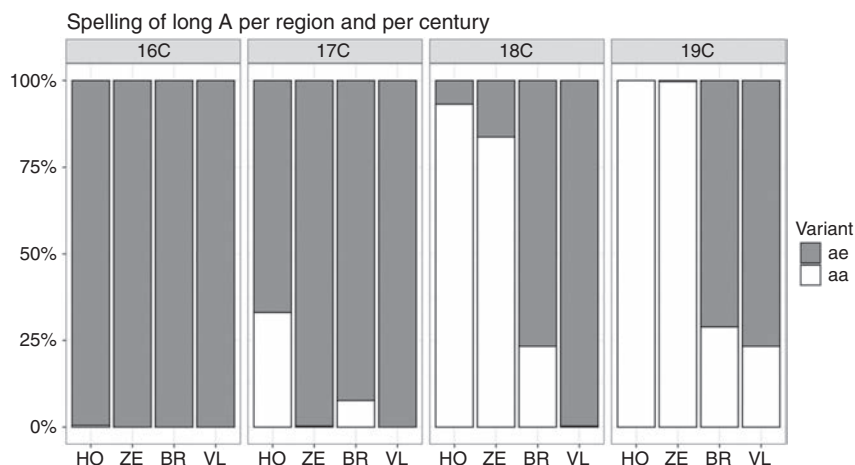


Figure 30.1 The incremental change of <ae> to <aa> across centuries



**Figure 30.2** The change from <ae> to <aa> across regions and centuries

as the first stages to a traditional S-curve, seems to suggest that spread could be taking place across generations.

However, splitting up the data per region, as in Figure 30.2, we see how the rise of <aa> rather seems to be due to geographical diffusion, originating from the northern center of Holland, and spreading across regions from there. In the mid-seventeenth century, Holland clearly leads the change, but by the mid-eighteenth century, the innovation has been diffused from Holland across the north, to the more peripheral region of Zeeland. At the same time, it seems to have also spread to the south of the language area, gaining ground in the more central region of Brabant. It is not until the nineteenth century, however, that it has finally spread to the more peripheral region of the south, Flanders, at a time when the change was already completed in both the center and the periphery of the north. As such, we can see a shift in the pluricentric landscape of the Dutch-speaking Low Countries, which changes from a constellation where Holland is the normative center where this innovation arises in the seventeenth century and from where it subsequently spreads, to more of a strict division between the north and the south of the language area from the eighteenth century onward, where we can also observe Brabant as an additional normative center for the southern territories. The eighteenth and nineteenth centuries show fairly distinct and thus presumably separate dynamics in the north and the south, precisely at the time when the political and sociocultural border between both areas started to develop and grow in importance.

These results illustrate how the spread of the incoming orthographic <aa> variant is a case of diffusion across regions, while also offering significant diachronic depth so that we can link the findings to the ongoing standardization process. However, rather than limiting ourselves to a description of how <aa>

won out by establishing itself in the dominant region of Holland, and was thus taken as input for further selection in the seventeenth century, we argue that this variable contributed to the development of a pluricentric normative landscape, with different centers (in north and south) from which innovations spread: in the eighteenth and nineteenth centuries, <ae> was clearly the dominant form in the south. By investigating diffusion processes at such a macro level across time and space, we believe it is possible to avoid the pitfall of a teleological view of standardization, thus contributing to a more sociolinguistically informed approach to orthographic history.

### 30.6 The Applicability of Transmission and Diffusion in Diachronic Orthography

We have discussed the concepts of transmission and diffusion, departing from their definitions in the original Labovian sense, with a focus on the type of language learning or acquisition involved, as well as in the Milrobian sense, which emphasizes the type of contact among speakers, either between or within communities. We have also discussed particularities of orthographic change, mainly centered around the observation that spelling is by definition a learned behavior. As such, we have argued that diffusion can be seen as the main source of orthographic change, which can typically be described as change ‘from above’ in the Labovian sense. In fact, we have observed that changes often spread across geographically defined communities in a process which can also be labeled supralocalization: regionally bound orthographic practices in the late medieval period thus often develop into supraregional writing traditions in postmedieval times. Needless to say, such supraregionalization processes can provide fertile ground for later standardization efforts, although they often also lead to pluricentric linguistic landscapes, where various normative centers coexist at a national and/or at a regional level.

Another topic addressed in the chapter has been the possibility of transmission in orthographic terms, as many changes in spelling practices actually seem to follow normal S-curve patterns with relatively incremental increases in usage. From this perspective, we can foreground the fact that orthography is also often picked up in a less conscious manner, for instance through reading – in that sense, orthography is not so different from other levels of the language. Therefore, we have argued, transmission can also be a useful concept in particular historical settings, for instance where we can link continuity in orthographic conventions to instruction and teaching practices, or in cases of close-knit networks and dense community structures, where changes are transmitted from one user to the next. The distinction between transmission and diffusion, however, can only be deemed useful if both concepts provide us with different models and different predictions concerning patterns of

change. We wonder, in other words, what we can gain by adopting the concept of transmission, and whether we cannot merely limit ourselves to diffusion as the main driver of orthographic changes. The view we have taken in this chapter is that the Labovian distinction in terms of learnability, based on the opposition of child versus adult language acquisition or learning, is harder to apply in historical settings, as cases of transmission, according to this perspective, demand fairly extensive investigations of the exact mechanisms underlying transmission. We deem the Milrobian distinction between local and supralocal changes to be more useful, which can be linked to transmission and diffusion respectively by foregrounding the scale of the contact situation.

Building on the contact and community-based perspective of Milroy (2007), we have also observed that diffusion-type changes in cases of spelling seem to follow more gradual, incremental patterns of change over the centuries, in line with the traditional S-curve model. This was, among other examples, the case for the spread of the innovative <aa> spellings in Dutch, as discussed in the final case study: <aa> forms gradually appear and spread across time and place between the seventeenth and the nineteenth centuries. At the same time, however, cases of transmission of spelling norms do allow for more abrupt and possibly also deliberate changes, as we have discussed in the case of the Dutch spelling and educational policy, which caused significant changes in orthographic practices in a relatively short period of time (Krogull 2018). In addition, note that we have treated both concepts mostly as independent concepts, and we have not focused much on the interaction between transmission and diffusion. In reality, of course, both types of changes co-occur and indeed interact. In Section 30.4.1, for instance, we discussed the regional diffusion of orthographic practices indicating a long *e* merger from Holland to Zeeland, but in fact, this change ‘from below’ was later reversed to a significant extent by a change ‘from above’. This process occurred when the officialized Siegenbeek norms (see Section 30.3.1) reintroduced the distinction between both etymologically distinct *e* sounds, and such norms were transmitted within the community through education (Rutten et al. 2020). Such and similar instances, where transmission overrules diffusion, open up promising avenues for future research, but fall outside of the scope of the present contribution.

### 30.7 Conclusion

Taking the observations and analyses in this chapter together, we argue that orthographic changes resemble changes in other levels of the language in the sense that they all allow for changes of both the local, transmission type, and the supralocal, diffusion type. They also differ, however, in the sense that

orthographic changes seem to be more incremental in cases of diffusion, and possibly more abrupt in cases of transmission – which is exactly the opposite from what you would expect for other domains of language. Taking a diachronic perspective on spelling variation and change, we observe gradual geographical, social and other types of diffusion as fairly typical. Overall, a highly frequent and therefore important pattern of change seems to be that of a transition from local or regionally circumscribed writing traditions to the supralocal level by means of (geographical) diffusion, often followed by standardization on the one hand, but also leading to situations of pluricentricity on the other hand.