From status to corpus: Codification and implementation of spelling norms in Luxembourgish[1]

Peter Gilles (Luxembourg)

1. Introduction: The status change of Luxembourgish

This chapter is concerned with reconstructing the processes of codification and implementation of orthography by exploring how certain norms find their way into actual language use. The test case chosen here is the codification and subsequent implementation of orthography in Luxembourgish. As a primarily spoken language, originating from a Central Franconian dialect and achieving high positive prestige as the national language since the end of the 19th century, Luxembourgish can today be regarded as the an *Ausbau* language, which has gained access into several new domains. Although historically closely related to German, Luxembourgish is today evaluated by the speech community as a language of its own. Learned as the first language of the majority of Luxembourgers, the language is embedded into a complex multilingual setting where it shares domains and functions with German and French (cf. Horner & Weber 2008; Fehlen 2009). (Standard) German is employed as the language of alphabetization in primary schools, as the language of media consumption and in general as a written language. French has the status of a written language and is also the most important lingua franca in the public space and at many workplaces. Because of a high proportion of foreign residents due to immigration (44%), the language situation is characterized by extensive multilingualism with a predominance of French (cf. Fehlen 2011). While both German and French are intensively taught in school, Luxembourgish, although the first language of approximately 60 per cent of the population, forms part of the school curriculum only rudimentarily. Furthermore, the government is rather reluctant to give Luxembourgish a more prominent role in the educational system. Despite this lack of active official support for Luxembourgish, since the 1970s one can observe a constant increase of usage in various domains, both as a spoken and written language. The reasons for this increase lie in the strong ideological link between identity and Luxembourgish and in the tremendous increase of informal literacy in the so-called new media where

Luxembourgish is used quite naturally (cf. Gilles 2011). One further reinforcing factor is the growing number of people acquiring Luxembourgish as a second or third language. According to the recent census of 2011, for the total of approximately 500.000 inhabitants 56 per cent reported that Luxembourgish is their 'best' language, while 71 per cent indicate that they use Luxembourgish at home, in school or at the workplace (cf. Fehlen et al. 2013a; Fehlen et al. 2013b).

In this context language standardization is taking place and has reached a medium level (cf. Gilles & Moulin 2003): While – in the terms of Haugen's (1966) model for standardization – norm selection has taken place by selecting the Central Luxembourgish variety as some kind of spoken standard (cf. Gilles 2006), the subsequent steps of norm codification, acceptance/implementation and elaboration are partially still in their incipient stages. Standardization is reflected in a remarkable increase in norm awareness in the society, as can be deduced from the following observations. The idea that a norm exists is gaining a foothold in public and private debates – although people probably have only partial knowledge of the concrete norms themselves. When Luxembourgish is used by official institutions on public signs or texts, in advertising campaigns of big companies, on election posters etc., one will only find correct Luxembourgish. In formal letters or emails it is common to apologize for probably not writing correctly.[2] Also observable is an increasing demand for language correcting material (dictionaries, course books, orthography guides etc.) or software. Worthy of mention is the orthography correction software *spellchecker.lu*.[3] This website provides a convenient possibility to check and correct one's own texts for orthographical mistakes. According to the developer, an average of 1000 people per day use the software. Between 2009 and 2011 approximately 39 million words were corrected by the system.

In this context of incipient norm codification, norm implementation and norm awareness the crucial research question is: How do (orthographical) norms enter language use, although language education is not supportive of them at all? Put differently in standardization terms: How are changes in the language status reflected in the language corpus?[4]

This chapter is organized as follows: Section 2 will introduce the top-down perspective of codification by giving insight into the historical development of the official orthographic regimes for Luxembourgish. Section 3 then is dedicated to implementation aspects of the official orthography. After presenting the by and large

emerging text genres that adhere to the official rules and which thus can be conceived as 'model texts' (literary texts, mass media text production etc.), a corpus-analytic study will be presented to trace actual compliance with the official orthographic rules.

2. Official orthographic regimes for Luxembourgish

In this section a concise history of norm codification, the second step in Haugen's (1966) standardization theory, will be given. I will concentrate on the language-ideological motivations behind the choice of a certain orthographic regime.

2.1 The Welter-Engelmann system (1914)

The beginning of text production in Luxembourgish in the 19th century was characterized by literary attempts of individual authors who were struggling between rendering the phonetics of their dialect as accurately as possible and adapting orthographic features from German and (to a far lesser extent) French (see Newton 2000).[5] Efforts towards standardization were not discernible and it was not before the beginning of the 20th century that the need for some kind of standardized form of Luxembourgish (which was still regarded as a dialect) was felt. This first system, called the Welter-Engelmann system, was developed by the historical linguist René Engelmann for the first schoolbook of Luxembourgish texts. The reader with the title 'Das Luxemburgische und sein Schrifttum' (Welter 1914) was an immediate reflex of a new educational law from 1912, where Luxembourgish was introduced as a minor subject in schools for the first time. Although the system had only a quasi-official status, it was broadly accepted was taught in secondary schools nationwide and for many years until the 1950s. Even today one can find sometimes remnants in the unmonitored writing of elderly people.

The rules were primarily conceived for reading Luxembourgish literature and enabling the pupils to get acquainted with the first literary texts of the 19th century. Writing Luxembourgish was foreseen only to a far lesser extent. With Luxembourgish still regarded as a dialect of German, it was also not intended to promote Luxembourgish for further domains – because German and French were considered the important written languages. The dependency on German and also German spelling rules are explicitly stated in the introduction where it says that the Luxembourgish dialect stems from German and that relying on the German rules will help the pupil, as he or she is learning German anyway.

[…] Anlehnung an die deutschen Rechtschreibungsregeln, was umso eher gestattet war, als unsere Mundart germanischen Stammes ist. Damit ergibt sich vielfach eine treue Wiedergabe des deutschen Wortbildes, wodurch der kleine, noch ungeübte Schüler vor mancher überflüssigen Schwierigkeit und störenden Verwirrung geschützt bleibt.[6] (Welter 191, p.4)

The rules (written in German) themselves cover roughly only four pages and contain frequent remarks like "wie im Deutschen" ('like in German') or "der dem Deutschen fremd ist" ('foreign to German'), underlining that German is seen as kind of matrix for the Luxembourgish dialect. Given that Luxembourgish has developed phonologically away from German unnecessarily complex rules are introduced, for example:

Endkonsonanten werden geschrieben, wie sie gesprochen werden; also scharf, wenn nicht dem luxemburgischen Wort ein gleichlautendes im Hochdeutschen mit weichem Endkonsonanten zur Seite steht.[7] (Welter 1914, p.138)

Thus, the adjective *midd* 'tired' is written with <d> because of the German form *müde*, whereas *gutt* 'good' is written with <t> because of the German form *gut*, although both words end with a final voiceless [t]. This example demonstrates the mixture of different approaches to orthography: Language historical aspects are mixed with phonetic ones. This orthographic principle is still in effect in the present-day rules and remains a major source of spelling mistakes.

To summarize, this first spelling system largely depends on the German system, although the phonetics of Luxembourgish are not represented well. It seems that the – at least felt – language historical relationship with Standard German was impeding the development of an own and more independent orthographic system.

2.2 The Margue-Feltes system (1946)

Immediately after World War II the pressure was high in the government and especially the educational sector to diminish the cultural relationship with Germany and the German language after the cruel occupation by Nazi Germany. Already in

1946 the so-called "Margue-Feltes" system was officially proclaimed. The pronounced ideological stance is mentioned at the beginning of the ministerial decree: Firstly, the need for a uniform orthography for the national language is clearly expressed. Secondly, Luxembourgish should be taught in schools and therefore the pupils would need precise spelling rules.

> Considérant que l´emploi de notre parler national comme langue écrite se heurte surtout à l´absence d´une orthographe uniforme ;
> Considérant que si la langue luxembourgeoise doit former matière d´enseignement, il est indispensable de donner aux élèves des règles nettes et précises d´après lesquelles s´écriront les mots de la langue.[8] (Mémorial, 1946)

Thus, this orthography conveys not only spelling rules, as one would expect as a codificational task, but also expresses an explicit language policy on how to implement Luxembourgish as a written language.

In order to avoid the dependence on German orthographical principles as far as possible, the phonetician Jean Feltes devised a strict phonetically based orthography. Thus, letters and letter combinations obviously coming from the German system like <ä>, <ß>, <sch>, <ck>, <ie> <tz>, the length indicator <h> and others were all eliminated. By assigning one sound to exactly one letter and by capturing as many phonetic details as possible without using uncommon letters, the result was an astonishingly consistent but heavily complex, if not unusable, system.[9] With all the knowledge of the German spelling reforms, it seems inevitable that introducing a radical phonetic orthography could only lead to disaster. Still, from a sociolinguistic perspective this attempt also underlines how the attitudes towards the national language strongly called for the creation of a distinct and independent written language and for standardization in general. The Margue-Feltes system of 1946 was the first orthographical system officially issued by the Government and was foreseen for teaching in schools as well. But the break with tradition, triggered through an ideological move away from everything that was related to German, was far too radical. When writing or reading Luxembourgish at all, people were used to a system close to German and they categorically refused the new system. Moreover, the rules were soon trivialized, ridiculed and discredited by several articles in the written press. This spelling reform, therefore, was a nightmarish catastrophe. Although not officially

withdrawn, the system was hardly ever implemented thoroughly and was buried in the files.

2.3 Orthography of the *Luxemburger Wörterbuch* (1975)

In the context of the multiannual project of the *Luxemburger Wörterbuch* (LWB, 1950-1977) it was deemed necessary to develop a new and simplified spelling system. The main motive this time was a scientific one: to create the dictionary, which in the optimal case should also be easy to read. For the lexicographic work, analysis of language history was necessary and by doing so the resemblances to the German spelling system as well as the old Welter-Engelmann system of 1914 came back into the game. In the introduction to the *Luxemburger Wörterbuch* it is stated that the rules are formulated with regard to tradition and by respecting the German or French structure of the word. In general, the spelling is a rather phonetic one for the vowels and it is oriented more to the German system with regard to the consonants. Also borrowed from German is the capitalization of the first letter in nouns, which is then applied to French loans as well.

Compared to the unsuccessful 1946 attempt, this new orthography could be considered as a step backwards: Firstly, the spelling system was originally only intended for the limited purposes of the dictionary, for example for language conservation. Because of the many regional variants of hitherto largely unstandardized Luxembourgish, the dictionary makers were faced with the challenge of including as many variants as possible. The main driving force was thus language documentation. Unsurprisingly, therefore, no standardization or language planning actions were foreseen to propagate these rules for a broader public or for further domains of writing. Note that in these times writing in Luxembourgish was still rather uncommon anyway. Nevertheless, it is noteworthy that the lexicographers showed no pronounced interest in promoting the spelling system on a broader scale. Secondly, in various text passages in the dictionary the authors showed a peculiar indecisiveness as to whether Luxembourgish was regarded as a dialect of German or a language on its own. Nevertheless, the rules were a pragmatic compromise between different forces and after the completion of the *Luxemburger Wörterbuch* the dictionary rules were

made official in a ministerial decree in 1975 (Mémorial, 1976). With a slight reform in 1999 (Mémorial, 1999), they provide the basis until today.

2.4 Present-day spelling commission (2009 onward)

With the beginning of the 21st century, profound changes in the speech community again call for a reconsideration of the spelling issue. Several factors have to be mentioned here: (1) The self-conception of Luxembourgish has strongly moved away from the notion of being a dialect of German, this development cumulating in the language law of 1984. In this context, writing correct Luxembourgish is increasingly becoming a desirable skill. (2) More and more people from the 44 per cent share of foreign residents are starting to learn Luxembourgish, either informally as a second or third language as children or teenagers or in language classes for adults. (3) The increasing importance of internet-related media in the everyday-life creates new written domains and also new text genres.

By applying the terms of von Polenz, we are in the middle of the stages of what he has termed the "popularization" and "pedagogization" of language norms (cf. von Polenz, 1999, p.3) which call for a further reformulation of the orthography, that is clarification and careful simplification. Thus, in 2009 a new working group was officially appointed by the Ministry of Culture with the task to modify and simplify the current rule system. Members from the Ministry of Culture, the Ministry of Education and the University, representing different skills and interests, constitute this working group. The new rules will also be applied in the newest dictionary of Luxembourgish, the *Lëtzebuerger Online Dictionnaire* (LOD, 2007ff.).

To summarise the history of orthographic standardization, one firstly has to note the overt preoccupation with issues of Luxembourgish orthography despite Luxembourgish having the status of a predominantly oral language. More importantly, one can observe how the (socio-)linguistic evolution of Luxembourgish, that is *Ausbau*, in the context of its multilingual environment is mirrored in the rationale for the standardization of the orthography. At the beginning of the 20th century Luxembourgish was regarded as a dialect of German with only little usage as a written language and hence the orthography was oriented strongly towards the German model. The post war reform of 1946 had a pronounced anti-German

underpinning, resulting in a radically new version without any resemblances to the German language. While the 1975 reform brought a return to a partial orientation towards German (and to lesser extent French), the currently planned reformulation of the orthographical system is owed to an increasing demand for norm-adhering text production.

In terms of Haugen's theory of standardization, this section has introduced the key aspects of "norm selection" and "norm codification" (cf. Haugen, 1966). The following section will be devoted to questions of the implementation of these norms in actual writing in the speech community.

3. Bottom-up approaches to Luxembourgish spelling

As mentioned in the introduction, the present-day speech community has seen a general increase of written use of Luxembourgish across nearly all domains since the 1990s, mainly due to the factors of identity construction and the ease of use of written language in digital media (e-mail, SMS, social networks, guest books, personal homepages etc.) (cf. Gilles, 2009; 2011). On the other hand, due to an educational setting in which Luxembourgish is taught on a rudimentary level, if at all, orthographic skills are distributed unevenly in society. We find a (small) group of language professionals including (literary) authors, translators, journalists, teachers, language activists, who apply the norms in their text production and who can be seen as "model writers" (cf. Ammon 2003) for other writers. On the other hand there is a large share of language users without training in orthography at all who nevertheless produce large amounts of Luxembourgish texts mainly for private purposes. This unbalanced distribution of skills is of course reflected in text production and to capture this field of different "orthographic regimes" Mark Sebba's concept of the "orthographic space" is a valuable heuristic for further analysis (cf. Sebba 2009, p.41ff.). In the core of this space one finds what he calls the "fully regulated space" of "published texts and "'school' writing" with a complete adherence to standard norms. Further away from this core are text genres with a less regulated or even unregulated orthographic regime, among them 'chat room 'talk'', 'personal notes', 'personal letters', 'personal e-mails', 'SMS text messages', 'poetry', 'advertising', 'fanzines' or 'graffiti'. To varying degrees orthographical norms deviate in a more or less structured way from the fully regulated regime and these deviations can be analysed as having social meaning.

In the remainder of this paper, I will explore the orthographic space of Luxembourgish written texts, briefly starting out with the fully regulated space before presenting a corpus analytic study for the less regulated space.

3.1 Examples from the fully regulated orthographic space

Beginning in the 1980s, literary text production and translation constituted the predominant genre of published texts (cf. Berg, 2006) and these texts stick closely to

the official rules. A so far non-existing copy-editing has been implemented in the publishing process to ensure the correctness of spelling. Hence, forming part of the standardization process, one can observe the constitution of a fully regulated orthographic space. This includes also the translation activity of texts with a high symbolic value. For this context the translation of the gospel (Evangeliar, 2009) and the bible (currently in development) have to be mentioned and may serve as a reference point for correctly implemented orthography. Furthermore, the reports of parliament debates and of municipal councils are carefully drafted (including slight stylistic correction and the reduction of dialect variants) in correct spelling, as well as promotional texts of political parties.

A further example concerns the ambitious project of creating a comprehensive Luxembourgish version of Wikipedia (http://lb.wikipedia.org). Besides the Wikipediae in the big languages like English, French and German, numerous Wikipediae in smaller languages exist. The Luxembourgish one is created and maintained by a small group of enthusiasts, probably less than 50. At the moment it contains 38,000, often rather small, articles – compared to 19,000 in 2008 and 6,000 in 2006, thus showing the increasing interest in pursuing this ambitious goal. As one would expect from such a reference work, the authors attach great importance to the linguistic correctness of the articles. More relevant for our context here is that we can observe *in situ* how the official norms are implemented. First, on the discussion pages one can often find discussions about writing and about the criteria for selecting the most appropriate variant out of a group of alternatives. The most important goals are homogenization and the reduction of variants. Secondly, articles are often revised with all revisions still accessible in on the webpage, which then allow for the reconstruction of the writing process of an article (cf. Gilles, 2011). Without going into details here, it is obvious that these revisions form a valuable testing field for the implementation of orthographical norms. In doing so, the authors are testing and expanding the capabilities of Luxembourgish in various specialized domains. Although the Wikipedia articles serve the function of being 'model texts' written by 'model writers', it remains, however, unclear if not questionable, whether they will contribute to a further acceptance and implementation of language norms, as the readership for these small language Wikipediae is supposedly rather small.

With regard to public reach and acceptance and size, the most important source for written text production and reception is undoubtedly the Internet based

news portal of the local broadcasting company RTL (http://rtl.lu). This original branch of the RTL holding started as a local radio program in the 1930s and continues to be the most heard Luxembourgish-speaking radio station (cf. Newton, 2013). Alongside a TV channel (since the 1970s), RTL has developed a comprehensive internet platform for news since the 1990s. During incipient times, the format of the website was rather rudimentary and compliance with orthographical rules was low too.[10] Today, the website has developed into a fully-fledged and professionally maintained news portal, offering several types of news and related multimedia services through interaction with the radio and TV branch (see Figure 3.1).

Figure 3.1   Example page of the homepage of the news portal RTL.lu (27.8.2012)

According to the editors, some 60,000 users visit the site every day, with a steep increase in reach and acceptance over the last years. In relation to an approximate number of 400,000 people speaking Luxembourgish, this is an enormous figure. The editorial crew publishes between 60 and 100 articles per day, making the portal an ever-increasing source of written Luxembourgish. It goes without saying that nearly everything is orthographically correct. The authors and journalists have strict internal rules for ensuring orthographical and stylistic quality. The above-mentioned

spellchecking software has been integrated into the editing system. Compared with the incipient times, it is obvious that applying language norms has become a desirable goal for the RTL news company.[11]

Furthermore, the reader has several options to interact with the news portal, either by 'letters to the editor' or by commenting on articles. This latter option is widely used and news articles often receive between 150 and 200 comments from the readership. Through this participative structure, the news portal is establishing a public online community (cf. Androutsopoulos, 2003). The textual model given by the news article itself and by the numerous comments of readers, is encouraging more and more people to write Luxembourgish, despite their probably lacking spelling skills. Thanks to this news portal the Luxembourgish population is massively exposed to texts complying with the official norms. As a consequence, it is concluded that this exposure to texts conforming to the standard will enhance the competence of the reader through informal learning.

In concluding this section on the fully regulated orthographic space, it should have become obvious that this domain exists despite the fact that Luxembourgish is mainly used as a spoken language. Besides the 'classical' domains of the production of literary books and translation activities, Luxembourgish has gained ground in several public areas. It comes as no surprise that the mass media of today with their interaction between spoken (radio, TV) and written domains (web-based portals) represent a key factor in the tacit implementation of (not only) orthographic norms. It must be stressed here again that these developments take place without formal or otherwise supportive language policy actions of public administrative bodies.

3.2 A corpus-linguistic study of the less regulated/unregulated orthographic space

The less regulated and unregulated orthographic space probably represents the domain with the quantitatively highest usage of written Luxembourgish. This includes the private domain and the semi-public domain of the interactive mass media and social networks. This development is characteristic not only of the Luxembourgish situation but holds true for most of the speech communities influenced by the so-called new media: "more people write, people write more, and unregimented writing

goes public" (Androutsopoulos 2011, p.311). With regard to orthography, it can be assumed that 'unlicensed spellings' (cf. Sebba, 2007; Shortis, 2009; Tagg, 2009) are used to a certain extent, according to the needs of various practices, genres and styles. Thus, writing in a non-standard way on Facebook, on Twitter, in chatrooms or in comments to the news site RTL.lu does not imply per se that orthographic rules are unknown to the writer. Rather, the chosen orthographical deviations may serve certain social functions.[12]

This study, however, will concentrate only on one specific aspect, that is the compliance with the official orthographic rules on an overall scale. In a corpus-linguistic study the quantitative distribution of the norm-compliant and norm-deviant spelling forms will be examined, in order to evaluate the degree of implementation of these norms in the specific context of Luxembourgish not being taught formally in the educational system. To obtain reliable results it will be necessary to compile sufficient material for a corpus. The main criteria for the compilation of the corpus are: (1) Textual domains with widespread use in the society by many writers, (2) availability and accessibility of large data sources. These requirements are fulfilled for the Luxembourgish context only in the domain of the internet-based media as this is the domain where Luxembourgish is used most.

Two promising data sources for 'networked writing' (cf. Androutsopoulos 2011) could be identified: The first sub-corpus comes from public comments to news stories written by the readers of the news portal RTL.lu. The sizes of these reader comments range from one sentence to a whole page of text with an average of 68 words per comment. This sub-corpus contains all comments from the year 2012, i.e. a total of approximately 75,000 short texts written by a multitude of different writers.[13] The second sub-corpus is constituted by Internet Relay Chat data from the year 2004.[14] By having data from the years 2004 and 2012 the diachronic dimension can be utilized for the analysis. Note, however, that both text genres – in belonging to the less regulated/unregulated orthographic space – contain spellings of intentional, unlicensed deviations from the norm (e.g., for reasons of creativity or playfulness, especially in the youth language oriented chat room data) and which cannot easily be attributed to the writer's (lack of) competence with regard to the official norms. As the main idea of this corpus-analytic study is to estimate orthographical correctness on an overall scale and within a large corpus, it seems justified to disregard this variation. Both sub-corpora with their sizes of more than five million tokens each are

sufficiently large for a corpus-analytic study. The following table summarizes the main characteristics of the corpus.

Table 3.1   Main characteristics of the two sub-corpora

Corpus	Type of data	Tokens	Types
chat2004	Chat room interaction	5829483	229699
rtl2012	Public comments to news stories	5116143	212319

The running text of the comments and the chat data has been tokenized to obtain a list with all tokens of the texts combined with the word frequencies for the years 2004 and 2012.

One crucial methodological point concerns the measurement of compliance with the orthographical forms. In the simplest account, word frequencies of selected words are compared with their most common deviant forms. This approach is applied for a set of high frequency words in section 3.2.1. In order to obtain a more systematic overview it is necessary, though, to perform calculations for the compliance with certain orthographic rules or orthographic phenomena (for example, the spelling for diphthongs or the marking of length in vowels). To achieve this, one needs comprehensive lists with the correct word form and – more important – with the possible spelling mistake(s). Given the hardly existing training in Luxembourgish spelling through the educational system, spelling may indeed vary greatly, making a list of the most frequent possible mistakes crucial for analysis. Fortunately, such a list does exist in form of the correction list from the widely used online spellchecking system *spellchecker.lu*, which has been presented above. From 2009 to 2012 a total of 1,545.042 misspelled words (tokens) have been corrected, with these tokens relating to 18,4536 different misspelled words (types, including inflected word forms).[15] An extract of this list is shown in Table 3.2.

Table 3.2   Extract of the correction list from *spellchecker.lu*

Deviation	Correct	Gloss	Count	Rank
hun	hunn	'have'	43763	1
sin	sinn	'are'	27908	2

dei	déi	'this/these'	26592	3
gin	ginn	'give/go/become'	25158	4
daat	dat	'that'	21245	5
…				

This structured list, then, serves as the backdrop for the corpus analysis: Combinations of the deviant form and the correct form are quantitatively determined in the two sub-corpora. Its size of more than 180.000 word forms is sufficient to cover most occurrences in the text corpus. All quantitative examination was undertaken using the statistical programming language R (R Core Team, 2013).

3.2.1 High frequency words

From the most commonly misspelled list those words have been selected that were misspelled at least 1000 times (deviant spellings representing potential homonyms have been excluded). These 33 high frequency words representing approximately one fourth of all the tokens have then been correlated with the corpus data. The results in Table 3.3 present the total numbers for the two years along with the percentage of correct forms. The data has been ordered according to the decreasing percentage in the rtl2012 corpus. As can be seen, the correctness values vary greatly. The average value for 2012 is 56 per cent compared to only 37 per cent for 2004. This increase by 19 per cent can very cautiously be taken as an indicator that the awareness of correct spelling is gaining ground in Luxembourgish. When comparing the difference between the 2004 and 2012 on a per-word basis (last column in Table 3.3) it is striking that for most words an increase in correctness is observable.

Table 3.3   Deviation and correct form for 33 high frequency words, N for the two sub-corpora, percentages of correct form, difference in percentage between 2004 and 2012

Deviation	Correct	N 2004	N 2012	corr. % 2004	corr. % 2012	diff.
fun	vun	15696	41931	89	99	+10

vier	vir	4803	5201	84	94	+10
ëtt/ët	et	49897	66194	100	94	-6
fier	fir	14500	44668	64	90	+26
matt	mat	34966	34808	96	86	-10
ierch	iech	4380	8682	53	83	+30
nett	net	85700	77595	99	81	-18
as	ass	91564	76788	66	79	+13
hudd	hutt	777	2844	25	73	+48
durch	duerch	1005	4624	65	71	+6
lang	laang	4293	5258	58	69	+11
Zeit	Zäit	1698	3460	9	56	+47
gangen	gaangen	1019	1120	30	56	+26
geschriwen	geschriwwen	1355	2187	38	56	+18
daat	dat	53626	66331	49	56	+7
iwer	iwwer	2245	8835	35	54	+19
hieren	hiren	158	2756	7	52	+45
waat	wat	34197	27510	48	51	+3
baal	bal	2531	2329	18	50	+32
daag	Dag	1303	3040	24	49	+25
naischt/neischt	näischt	2425	9173	10	46	+36
natiirlech	natierlech	677	2124	64	46	-18
Zeit/Zait	Zäit	2776	4656	6	42	+36
sin	sinn	49445	41008	8	40	+32
wees	weess	10610	3932	2	39	+37
gin	ginn	21756	37380	6	37	+31
gesaat	gesat	105	679	18	34	+16
wärt	wäert	277	2810	8	30	+22
laafen	lafen	736	823	9	30	+21
hun	hunn	34752	33973	4	29	+25
gut/gudd	gutt	18433	11378	5	29	+24
kaafen	kafen	618	1469	8	29	+21
nie	ni	3671	4306	15	23	+8

Average		37	56	+19

Without going into details of the specific orthographic issues for these words, it is worthwhile to discuss the development of the following four specific words. The auxiliaries *ginn* 'to go/to give/to become', *sinn* 'to be', *hunn* 'to have' and *ass* '(he/she/it) is' were subject to the spelling reform of 1999. Until then they were written with a single final consonant, that is, *gin, sin, hun, as*. These words have all increased massively since 2004 (*ginn* +31%, *sinn* +32%, *hunn* +25%, *ass* +18%), thereby demonstrating that the official norm is, for whatever reason, gaining ground quantitatively. Although their rates partly still reside below 50 per cent, the growth in correctness is conspicuous and has to be attributed to the rising attention people pay to orthographic correctness.

3.2.2 Orthographic rules and orthographic phenomena

A far more differentiated picture of the implementation arises when certain orthographic rules or orthographic phenomena are analysed. In doing so, certain areas with varying norm compliance can be identified. The following discussion will concentrate on known problems of Luxembourgish orthography.

Diacritics
The first orthographic rules concern the placement of diacritics like <ä>, <é> and <ë>, which play an important role in the orthographic system to differentiate the numerous half-open front vowels. The dataset is characterized by fast and largely unplanned writing. Depending on the keyboard that is used an extra key is needed to type a character with a diacritic (Luxembourgers use either Swiss-German, German or French keyboards). Especially in chat, writing fast is crucial and accented characters are rather likely to lose their diacritics. For both sub-corpora one also has to take into account that technical constraints of the keyboard used and/or keyboard layout have an impact on the use of accented characters. It is expected to find less of these characters. This tendency is clearly shown in Table 3.4. The letter <ä> is fairly well represented in the correct way in both datasets (84% and 92%, respectively) – probably due to the influence of German, where <ä> is deeply rooted in the

orthographic system. Also, the diphthong <éi> reaches a rather high value in 2012 (64%). All other letters with diacritics, though, including <ë> (representing a schwa-like sound in root syllables) as one of the most used letters, are only used correctly in below 50 per cent of cases.

Table 3.4   Orthographic correctness for characters with diacritics

Type	Example	N 2004	N 2012	corr. % 2004	corr. % 2012	diff.
*<ae> →[16] <ä>	*<Haenn> for <Hänn> ‚hands'	9021	32853	84	92	+8
*<ei> → <éi>	*<Sei> for <Séi> 'lake'	52283	71196	5	64	+59
*<e, ö> → <ë>	*<schlemm> for <schlëmm> 'bad'	111091	159939	0	39	+39
*<eck> →<éck>	*<Breck> for <Bréck> 'bridge'	7436	9303	7	35	+28
*<eng> →<éng>	*<drenken> for <drénken> 'to drink'	2447	11699	5	30	+25
*<ech> →<éch>	*<secher> for <sécher> 'safe/sure'	4863	11096	0	17	+17
*<eeen> →<eeën>	*<leeen> for <leeën> 'to lay'	673	772	0	22	+22
Average				14	43	+28

While taking into account that these low correctness rates are partly related to the text genre, the considerable increase from 2004 to 2012 is nevertheless remarkable and can be regarded as an incipient stage in the implementation of the official rules. One can also assume that the rates would even be higher if a more planned text genre (for example, emails, private letters) had been studied.

Table 3.5 contains the results for spelling variation for three of the eight diphthongs of Luxembourgish. For the phonetically rather short diphthongs <ie> and <ue> the offset is omitted sometimes, but in general the correctness rate is high and

clearly increasing from 2004 to 2012. The increase is especially high for the diphthong <äi>, whose correct spelling was virtually absent in 2004.

Table 3.5   Orthographic correctness for diphthongs

Type	Example	N 2004	N 2012	corr. % 2004	corr. % 2012	diff.
<u>,<o> →<ue>	*<uwen> for <uewen> ‚above'	17918	6825	96	94	-2
*<ur> →<uer>	*<durch> for <duerch> ‚through'	14389	14513	78	88	+10
<i>,<e> →<ie>	*<Bleder> for <Blieder> ‚leaves'	11744	13735	64	87	+23
*<ir> →<ier>	*<Dir> for <Dier> ‚door'	6328	15892	65	75	+10
*<ai> →<äi>	*<Sait> for <Säit> ‚page'	23664	35517	9	60	+51
*<ei> →<äi>	*<Seit> for <Säit> ‚page'	35814	42685	6	50	+44
Average				53	76	+23

The spelling of the diphthong <äi> still shows somewhat less correctness rates today compared to the others, which is partly due to the phonetic variation ranging from a closed variant [æˑɪ] to an open one [aˑɪ] (see Gilles & Trouvain, 2013). This variation is reflected in using a more 'open' variant <ai> next to the 'closed' one <äi>. The emergence of the variant <ei>, on the other hand, is related to the spelling of the cognate words in Standard German (compare, for example, Luxembourgish <bäissen> vs. Standard German <beißen> 'to bite'). Thus, when using a variant like *<Seit> the writer is not clearly distinguishing between German and Luxembourgish. For the diphthong <äi> one can thus observe a competition between variants originating from different sources.

The next three features in Table 3.6 also relate to Standard German. The first feature concerns the distribution of <f> and <v> (both representing the voiceless fricative [f]), which is the same in the two languages. Deviations concerning this feature occurred in 2004 but are nearly absent in 2012.

Table 3.6   Orthographic correctness for consonants

Type	Example	N 2004	N 2012	corr. % 2004	corr. % 2012	diff.
*<f> → <v>	*<fun> → <vun> ,of/from/by'	34164	87432	79	97	+18
*<esch> → <eg>,<ech>	*<bëllesch> → <bëlleg> ,cheap'	899	4581	92	99	+7
*<ch> → <sch>	*<peschen> → <pechen> ,to glue'	278633	258251	97	98	+1
Average				89	98	+9

Furthermore, like Central Franconian, Luxembourgish is also subject to coronalization of the palatal fricative [ç], which is fronted to [ɕ] and sometimes even to [ʃ] (see Newton, 1993). Note that this sound change is not (yet) captured in Luxembourgish spelling, instead <ch> or <g> are used like in German. The close phonetic merger may then give rise to deviations like *<esch> instead of <eg> or <ech>. In addition, the fricative [ʃ], written as <sch>, can in turn be mixed up with the phonetically close fricative [ɕ], giving rise to the deviation <ch>. As can be seen from Table 3.6, the correctness rate for both features is nearly 100 per cent. These high values are probably due to the importance of Standard German as the language of alphabetization in the school system and probably have less to do with an increasing acceptance of the orthography.

Due to sound change, the full vowel in suffixes with a former [i] was centralized to schwa, which is expressed in the orthography. Thus the Standard German suffixes <lich>, <isch>, <ig> appear as <lech>, <esch>, <eg> in Luxembourgish. The results in Table 3.7 clearly demonstrate that this feature is

correctly realized on a nearly absolute level. The few percent of <i> spellings can be attributed to German influence. Furthermore, there is only little increase between 2004 and 2012, indicating that this orthographic feature already has a stable representation in the corpus.

Table 3.7   Orthographic correctness of the vowel in suffixes

Type	*Example*	*N* *2004*	*N* *2012*	*corr. %* *2004*	*corr. %* *2012*	*diff.*
*<lich> → <lech>	*<wierklich> → <wierklech> ‚really'	899	4599	92	99	+7
*<isch> → <esch>	*<komisch> → <komesch> ‚strange'	1438	6596	94	96	+2
*<ig> → <eg>	*<richtig> → <richteg> ‚correct'	8607	13194	96	96	0
Average				94	97	+3

As part of connected speech processes, certain schwas can be elided even in a normal rate of speaking, especially in the frequent syllable sequence <éieren>, which is then often realized as [ɜiʁən] instead of [ɜiəʁən]; compare, for example *<bougéiren> vs. <bougéieren> 'to move'. According to the orthographic rules these reduced forms are not correct. Table 3.8 reveals that reduced forms were quite frequent in the 2004 chat corpus, indicating that writing according to the phonetic principle was the norm. For 2012, a tremendous increase of 56 per cent can be observed for the correct forms. (Note, however, that the number of tokens for 2004 is far less than 2012.)

Table 3.8   Orthographical correctness for the suffix <éieren>

Type	*Example*	*N* *2004*	*N* *2012*	*corr. %* *2004*	*corr. %* *2012*	*diff.*
*<éiren> → <éieren>	*<héiren> → <héieren> ‚to listen /to hear'	904	32568	25	81	+56

This feature was selected to exemplify how the natural spelling of a frequent phonetic process is dismantled in favour of the prescribed form. Note, that the phonetic process of schwa elision still is quite active, it is simply that the official norm is gaining ground. From the perspective of standardization, I regard this development as a consolidation and stabilization of the official norm.

The discussion of the orthographic rule governing the representation of vowel length will conclude this corpus analysis. In front of a single consonant letter, the vowel is interpreted as long (<molen> [mo:lən] 'to draw'), in front of two consonants, the vowel has to be geminated in order to be interpreted as long (<mools> [mo:ls] '(you) draw'). A single vowel before two or more consonant letters is regarded as short (<mussen> [musən] 'must'). Duplication of the vowel letter appears to be an iconic representation of vowel length but, as the examples illustrate, the rule also relies systematically on the number of consonants following the long vowel. This complex rule was first introduced in the 1914 spelling system and was modified with the 1975 spelling reform. It can be regarded as a characteristic of Luxembourgish spelling as it diverges greatly from the Standard German system. It goes without saying that in the given situation of a largely lacking training in Luxembourgish orthography, we expect the most deviations from the norm for this vowel length rule.

For the long vowels [o:] and [a:], the corpus analysis will be conducted from two perspectives for deviations, that is, (A) missing duplication of the vowel letter in front of two or more consonants and (B) redundant duplication in front of a single consonant. High frequency words like *<daat> or *<waat> are excluded. The results in Table 3.9 show that, except for the redundant duplication of *<oo>+C → <o>+C, vowel length is marked correctly only on an average level with values ranging between 48 per cent and 64 per cent in 2012. When compared to the results of the other features discussed so far, it can be concluded that the vowel length rule is not yet implemented on a considerably high level and due to the high level of variation a consistent norm has not yet been reached. However, compared to the 2004 corpus, an increase in correctness is still noticeable (average of +13%). The complexity of the rule and the lack of training are very likely responsible for these comparatively low values.

Table 3.9   Orthographical correctness for the marking of vowel length

Type	Example	N 2004	N 2012	corr. % 2004	corr. % 2012	diff.
(A) *<o>+CC → <oo>+CC	*<Stross> → <Strooss> ‚street'	7477	11738	44	61	+17
(B) *<oo>+C → <o>+C	*<Doot> → <Dot> ‚deed'	147634	130766	92	92	0
(A) *<a>+CC → <aa>+CC	*<Sach> → <Saach> ‚thing'	40793	60578	39	62	+23
(B) *<aa>+C → <a>+C	*<kaal> → <kal> ‚cold'	55728	75953	59	64	+5
(A) *<i>, *<ie>+CC → <ii>+CC	*<licht> → <liicht> ‚light'	1001	3899	47	48	+1
(A) *<u>+CC → <uu>+CC	*<Duscht> → <Duuscht> ‚thurst'	741	842	31	62	+31
Average				52	65	+13

After discussing several orthographic features, we finally conclude this section by calculating averages for the whole dataset. For a total of 33 orthographic features, most of them discussed here, the average correctness for 2004 is 51 per cent and for 2012 the average is 68 per cent, thus indicating an increase in overall correctness of 17 per cent in 8 years for the less regulated orthographic space represented by this dataset. Notwithstanding this, the results ought to be interpreted carefully, as they depend on the text genres and on the selection of the orthographic features. Of course not all possible features could have been analysed in this study. Also note that these statistics are based only on words that are contained in the list of frequent spelling errors in *spellchecker.lu*. This means that words written correctly were not included in the data analysis at all. Given the partly phonetic basis of the Luxembourgish orthography, it is easy to write several words correctly without knowing any rules. Thus, in the present context, it can be argued that the overall level of correctness is probably higher.

4. Conclusion

The starting point of this chapter was the on-going status change of Luxembourgish in the context of its multilingual situation, which manifests itself in an increased use of Luxembourgish as a written language (mainly in the digital media) and a growing norm awareness. These status-related developments are paralleled by the development of official spelling rules since the beginning of the 20th century. As regards norm codification, it is revealing to see how the basis for orthography is dependent on the socio-cultural and socio-political underpinning of the respective period of time. Here, the perceived status of Luxembourgish as either a dialect or as a language of its own and the (accepted or rejected) orientation towards the Standard German orthography have directly influenced the concept and design of the orthography. Today, norm codification of Luxembourgish spelling can be regarded as provisionally completed.

At first sight, norm implementation might be difficult to put into practice: Luxembourgish and its orthography is not taught in school as a proper subject and the transmission and diffusion of norms through the school system is hardly facilitated. Nevertheless, the orthographic norms are entering text production, starting with literary texts and translations (beginning in the 19th century, gaining ground since the 1980s). Unsurprisingly, the ongoing, early-stage implementation gives rise to different orthographic spaces. When Luxembourgish is used in the mass media, i.e. on news sites of TV or radio stations, it is noteworthy that the official orthographic rules are increasingly applied. Here, the well-known and heavily used news media portal RTL.lu plays a crucial role in offering large amounts of correctly written texts to the public. These texts can be characterized as 'model texts' (at least for orthography). A news portal like RTL.lu thus implicitly takes up the role of an important, non-official agent for norm implementation.

Next to these fully-regulated orthographical spaces, several less-regulated orthographical spaces have emerged, especially with the overwhelming use of Luxembourgish in the digital media. At the centre of the analysis is the corpus-analytic study of orthographical correctness in chat room data and comments to a news website. For the first time, corpus-driven results about the implementation of the official spelling in Luxembourgish are now available. The quantitative results show an overall correctness rate of 68 per cent for the most recent data set of 2012. When breaking down this overall average for various high frequency words and orthographic rules one realizes the varying degrees of correctness within the whole

orthographical system. Areas of far-reaching implementation (e.g., suffixes like <eg>, <lech>, <esch>, <éieren>) can then be identified besides areas of still incipient implementation (e.g., usage of <ë>, marking of vowel length). Using this method it has become feasible to quantitatively measure the very concrete implementation of spelling norms in texts. Considering the general lack of formal language training in Luxembourgish, this considerably high correctness rate is rather surprising and is clearly related to some kind of norm awareness and the (conscious or unconscious) intent to adhere to the official rules. This specific kind of norm implementation that is not related to top-down language planning actions I am calling 'tacit norm implementation': in this case codified norms seem to arise through various bottom-up processes and percolate into usage gradually. Direct imitation of other writers, which are accepted as model writers in the respective community, will play a major role in the process of adapting one's own spelling to a perceived correct one. This option seems to be obvious for the comments of readers to news stories presented above, because the reader comment is directly linked to the correct news story text. Of course, it remains to be seen how far this tacit norm implementation path can proceed as long as no sustainable official support by the education system is provided.

As regards to the method as such, improvements may be conceivable: the present approach predominantly relies on token frequencies and in a more refined examination it would be necessary to also incorporate type-based calculations to balance high frequency and low frequency words. Furthermore, the corpus analysis pooled the frequencies for all writers. A desirable next step would be a more fine-grained investigation on the basis of the individual writer in order to determine the individual degree of norm adherence. Independent of the method chosen, it has to be emphasized that if one takes the notion of norm implementation seriously, then quantitative, corpus-based studies similar to this one are needed.

Bibliography

Ammon, Ulrich. 2003. On the Social Forces that Determine what is Standard in a
   Language and on Conditions of Successful Implementation.
   Sociolinguistica 17. 1–10.

Androutsopoulos, Jannis. 2003. Online-Gemeinschaften und Sprachvariation.
   Soziolinguistische Perspektiven auf Sprache im Internet. *Zeitschrift für
   germanistische Linguistik* 31(2). 173–197.

Androutsopoulos, Jannis. 2011. Language change and digital media: A review of
   conceptions and evidence. In Tore Kristiansen & Nikolas Coupland (eds.),
   Standard languages and language standards in a changing Europe, 145–
   161. Oslo: Novus Press.

Armstrong, Nigel. 2013. *Standardization, ideology and linguistics.* Basingstoke ;
   New York: Palgrave Macmillan.

Berg, Guy. 2006. Abschied vom Dialekt. Zur lëtzebuergeschen belletristischen
   Gegenwartsliteratur. In Claudine Moulin & Damaris Nübling (eds.),
   *Perspektiven einer linguistischen Luxemburgistik. Studien zu Diachronie
   und Synchronie*, 341–356. (Germanistische Bibliothek 25). Heidelberg:
   Winter.

Evageliar. 2009. *Lëtzebuerger Evangeliar*. Luxembourg: éditions saint-paul.

Fehlen, Fernand. 2009. *BaleineBis : une enquête sur un marché linguistique
   multilingue en profonde mutation = Luxemburgs Sprachenmarkt im
   Wandel*. Luxembourg: SESOPI Centre intercommaunitaire.

Fehlen, Fernand. 2011. Le statut du français sur le marché linguistique du
   Luxembourg- Le choix de la langue comme enjeu d'un champ scientifiqie
   en devenir. In Peter Gilles & Melanie Wagner (eds.), *Linguistische und
   soziolinguistische Bausteine der Luxemburgistik*, 151–177. 1., Aufl.
   (Mikroglottika. Minority Language Studies 4). Frankfurt a.M. [et al]: Peter
   Lang.

Fehlen, Fernand, Andreas Heinz, François Peltier & Germain Thill. 2013a. *La
   langue principale, celle que l'on maîtrise le mieux / Die am besten
   beherrschte Sprache (Hauptsprache)*. (Recensement de La Population
   2011 - Prémieres Résultats). Luxembourg: STATEC.

http://www.statistiques.public.lu/fr/publications/series/rp2011/2013/
17-13-langues/index.html (accessed, 25 April, 2013).

Fehlen, Fernand, Andreas Heinz, François Peltier & Germain Thill. 2013b. *Les
langues parlées au travail, à l'école et/ou à la maison / Umgangssprachen*.
(Recensement de La Population 2011 - Prémieres Résultats).
Luxembourg: STATEC.
http://www.statistiques.public.lu/fr/publications/series/rp2011/2013/
13-13-langues/index.html (accessed, 25 April, 2013).

Gilles, Peter. 2006. Dialektausgleich im Luxemburgischen. In Claudine Moulin &
Damaris Nübling (eds.), *Perspektiven einer linguistischen Luxemburgistik.
Studien zu Diachronie und Synchronie*, 1–27. (Germanistische Bibliothek
25). Heidelberg: Winter.

Gilles, Peter. 2009. Jugendsprachliche Schriftlichkeit auf Luxemburgisch in den
Neuen Medien. In Charles Berg, Lucien Kerger, Nico Meisch & Marianne
Milmeister (eds.), *Savoirs et engagements. Hommage à Georges Wirtgen*,
166–175. Differdange: Éditions Phi.

Gilles, Peter. 2011. Mündlichkeit und Schriftlichkeit in der luxemburgischen
Sprachgemeinschaft. In Georg Mein & Heinz Sieburg (eds.), *Medien des
Wissens. Interdisziplinäre Aspekte von Medialität*, 43–64. Bielefeld:
transcript.

Gilles, Peter & Claudine Moulin. 2003. Luxembourgish. In Ana Deumert & Wim
Vandenbussche (eds.), *Germanic standardizations: past to present*, 303–
330. (Impact 18). Philadelphia: Benjamins.

Gilles, Peter & Jürgen Trouvain. 2013. Luxembourgish. *Journal of the
International Phonetic Association* 43(1). 67–74.

Haugen, Einar. 1966. Linguistics and language planning. In William Bright (ed.),
Sociolinguistics. Proceedings of the UCLA sociolinguistics conference, 1964,
50 – 71. The Hague, Paris: Mouton.

Horner, Kristine & Jean-Jacques Weber. 2008. The language situation in
Luxembourg. *Current Issues in Language Planning* 9(1). 69–128.

Jaffe, Alexandra, Jannis Androutsopoulos, Mark Sebba & Sally Johnson. 2012.
Orthography as Social Action: Scripts, Spelling, Identity and Power.
Berlin/New York: Walter de Gruyter.

LOD. 2007ff. Lëtzebuerger Online Dicionnaire - LOD. http://www.lod.lu/lod/
     (accessed 31 January, 2011).

LWB. 1950-1977. *Luxemburger Wörterbuch (5 Bände)*. Luxemburg.
     http://engelmann.uni.lu:8080/portal/wbb/woerterbuecher/lwb/wbgui.

Mémorial. 1946. Arrêté ministériel du 5 juin 1946 portant fixation d'un système
     d'orthographe luxembourgeoise. *Mémorial du Grand-Duché de
     Luxembourg* 40. 637–641.

Mémorial. 1976. Arrêté ministériel du 10 octobre 1975 portant réforme d'un
     système officiel d'orthographe luxembourgeoise. *Mémorial B - Journal
     Officiel du Grand-Duché de Luxembourg. Amtsblatt des Großgherzogtums
     Luxemburg. Recueil Administratif et Économique, B-Nr.68, 16 novembre
     1976* 68. 1365–1390.

Mémorial. 1999. Reglèment grand-ducal du 30 juillet 1999 portant réforme du
     système officiel d' orthographe luxembourgeoise. *Mémorial A. Journal
     Officiel du Grand-Duché de Luxembourg / Mémorial A. Amtsblatt des
     Grossherzogtums Luxemburg* 112. 2040–2048.

Newton, Gerald. 1993. Allophonic variation in Luxembourgish palatal and
     alveolar-palatal fricatives: Discussion of an areal survey taken in 1979. In
     John L. Flood, Paul Salmon, Olive Syce & Christopher Wells (eds.), *"Das
     unsichtbare Band der Sprache". Studies in German Language and Linguistic
     History in Memory of Leslie Seiffert*, 627–656. Stuttart: Akademischer
     Verlag.

Newton, Gerald. 2000. The spelling of Luxembourgish : Systems and
     developments since 1824. In Gerald Newton (ed.), *Essays on politics,
     language and society in Luxembourg*, 135–162. Lewiston: Edwin Mellon
     Press.

Newton, Gerald. 2013. Radio Luxembourg in Peace and War. *German Life and
     Letters* 66(1). 55–76. (accessed 5 February, 2013).

Polenz, Peter von. 1999. *Deutsche Sprachgeschichte vom Spätmittelalter bis zur
     Gegenwart - Band III: 19. und 20. Jahrhundert*. Berlin: de Gruyter.

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*.
     Vienna, Austria. http://www.R-project.org.

Sebba, Mark. 2007. *Spelling and society: the culture and politics of orthography around the world*. Cambridge: Cambridge Univ. Press.

Sebba, Mark. 2009. Sociolinguistic approaches to writing systems research. *Writing* 1(1).

Shortis, Tim. 2009. Revoicing TXT. Spelling, vernacular orthography and "unregimented writing." In Steve Wheeler (ed.), *Connected Minds, Emerging Cultures: Cybercultures in Online Learning*, 225–243. Charlotte, NC: Information Age Pub.

Street, Brian. 2003. What's "new" in New Literacy Studies? Critical approaches to literacy in theory and practice. *Current issues in comparative education* 5(2). 77–91.

Tagg, Caroline. 2009. A corpus linguistics study of SMS text messaging. Birmingham: University of Birmingham. http://etheses.bham.ac.uk/253/1/Tagg09PhD.pdf. (accessed 1 October, 2013)

Welter, Nikolaus. 1914. *Das Luxemburgische und sein Schrifttum*. 1. Auflage. Luxemburg: G. Soupert.

Endnotes

---

[1] I would like to thank Julia de Bres for valuable input to the present paper.

[2] See, e.g., the following extract from an email to the author: *Fir t'eischt well ech mech entschellegen dass ech heischtwarscheinlech keen richtegt Letzebuergescht schreiwen (ech sin et leider ni geleiert gin). [...]* 'First and foremost I would like to apologize for probably not writing correct Luxembourgish (It was never taught to me).'

[3] See http://spellchecker.lu. The first spellchecker was developed in the 1990ies (cf. http://cortina.lippmann.lu (last accessed 18.11.2013)). In the meantime Microsoft is also offering so-called 'language packs' for Luxembourgish, which offer a localised user interface and a spellchecking device for Office products (http://www.microsoft.com/lb-lu/download/details.aspx?id=6804 (accessed 18.11.2013)).

[4] Cf. also the similar case studies in Armstrong (2013).

[5] As for the writing system, Luxembourgish is, like French, written and printed since its beginnings in Roman type Antiqua. This is of particular interest as German, the genetically closest language to Luxembourgish, has been written in Luxembourg in Gothic print, like it was the case for all other German-speaking countries. Only two authors used Gothic print (André Duchscher (1840-1911) and Michel Rodange (1827-1876)), for which they were criticized by their peers.

[6] Translation: '[...] borrowings from the German orthographic rules, which are the more allowed, as our dialect originates from Germanic. This results in an often true rendering of the word, protecting the small and untrained pupil from unnecessary difficulty and annoying confusion.'

[7] Translation: 'Final consonants are written in the way they are pronounced, that is voiceless, if not the Luxembourgish word has a German cognate which is written with a voiced final consonant.'

[8] Translation: 'Considering that the usage of our national language as a written language is impeded because of the absence of a unified orthography and considering that the Luxembourgish language should be a subject in the educational system, it is indispensable that pupils have to be given clear and precise rules to write the words of the language.'

[9] An example: *Aaner lèuttipe gin nach flèicht fun dialèktshpézjaliste ssinjaleiert.* The same sentence in the present-day orthography: *Aner Lauttype ginn nach vläicht vun Dialektspezialiste signaléiert.* 'Other types of sounds will eventually be introduced by specialists for dialects.'

[10] Historic versions of the website are available through the 'Internet Archive Wayback Machine' (http://archive.org/web).

[11] I thank Luc Marteling, editor-in-chief at RTL.lu, for providing information on the usage of the news portal.

[12] Cf. the contributions in Jaffe et al. (2012) which illustrate in several case studies the social functions attached to orthography and to the deviations of orthography. Cf. also the approach of the New Literacy Studies (for example, Street 2003), where basically all literacy practices are regarded as ideologically influenced.

[13] For an illustrative example of these comments cf. for example, http://news.rtl.lu/wahlen/news/ 485423.html (accessed 10.11.2013).

[14] This data comes from log files of chat sessions from the no longer existing chat server 'www.luxusbuerg.lu' (1996-2011). I thank Caroline Döhmer for making this data available.

[16] The arrow → is used to indicate the mapping of the false form to the correct form.