Writing systems as modular objects: proposals for theory design in grapholinguistics

Abstract: This paper gives an outline of the Modular Theory of Writing Systems by answering the question: what are the elements or modules that are necessary for a writing system to work? A writing system is a notational system for a natural language. Based on this characterization, it is obvious that a necessary component of a writing system is a specific language system. What eventually constitutes a writing system in addition to this language system is a device that, put simply, relates units of a language system to units of a script. This component is termed ‘graphematics’ in the present framework and is regarded as a necessary module of a writing system. Above that, another typical component of writing systems, namely ‘systematic orthography’, applies to the ‘graphematic solution space’ and restricts the spelling possibilities of specific words in accordance to their belonging to a specific level of the vocabulary of the language. Supplemented by reflections on the status of scripts as well as of IPA as a writing system, an answer is finally given to the pertinent question how spoken language and written language are related to each other. The answer is that this relation is of a considerably indirect nature.

Keywords: writing system theory, grapholinguistics, graphematics, systematic orthography, conventional orthography, graphematic solution space, linguistic realism, script, phonetic transcription, International Phonetic Alphabet

1 Writing systems as abstract objects

This article aims to reflect on the question how a theory of writing systems should be designed. The notion ‘writing system’ is, in this case, restricted to notational systems of natural languages, to be distinguished from writing in general, which also encompasses notational devices for mathematics and music, etc. (cf. Harris 2009). The following considerations give a theoretical framing and an enhancement of the writing system theory that I have first published in Neef (2005) and that has evolved more elaborately in the past years. This general framework may be addressed as the Modular Theory of Writing Systems.

The first and most crucial question in order to devise a theory of writing systems is what kind of entity is a writing system. In my view, a writing system has the epistemological status of an abstract object. This characterization is akin to a conception of language systems as abstract objects, the crucial position of a scientific paradigm in linguistics that betimes is called Linguistic Realism (cf. Katz 1981, Neef 2014). Such an abstract object is what we relate to when we use it in speaking or hearing (i.e. in communication), in thinking, or in writing grammar textbooks. In contrast, knowledge of language (including the intermediate states of learning and losing a language) is something categorically different from the system proper. The scientifically relevant
objects of investigation ‘use of language’ and ‘knowledge of language’ call for fundamentally different theoretical approaches than the language system as an abstract object. The same holds for written language as well. The results of the use of both the abstract objects language system and writing system – i.e. spoken language and written texts – are empirical in nature in the sense that they can be observed directly. The systems that we reasonably assume to lie behind their use are of a different nature; they cannot be observed by any scientific method. The only way to make abstract objects palpable is to capture them in a theoretic model. As Eisenberg (1983: 43) puts it, a writing system can only be ‘reconstructed’. In this sense, language systems as well as writing systems are non-empirical objects. This is not to deny that writing systems have an existence in reality independent of the question whether or not they are captured in a theoretic model.

Realist approaches in linguistics are, of course, axiomatic in nature; they rest on basic assumptions that are at most plausible, but inherently non-provable starting points of reasoning. A specific linguistic theory is a valuable one if it enhances our understanding of the systems it aims to reconstruct. At the core of linguistic theories are definitions of concepts. For the present study, a central term is ‘writing system’. While many linguists take the terms ‘writing system’ and ‘orthography’ to mean more or less the same (cf. e.g. Sampson 1985: 19-20, Neef, Neijt, Sproat 2002: 1), I now take these terms to denote strictly different concepts (cf. Neef 2005, 2012), in line with a well-established tradition in German linguistics, going back at least to Eisenberg (1983). A writing system is a notational system for a natural language. It may or may not comprise a component of orthography. This idea will be unfolded later on in the course of argumentation.

2 The first module: a language system theory

In order to develop a theory of writing systems as abstract objects, different modules have to be specified. To this end, information that for the purpose of the model is regarded as given information needs to be distinguished from information that is affected by the mechanisms specific for a writing system. In a derivational conception of linguistic theory, this distinction is captured in the concepts of input and output. In a declarative conception of linguistic theory – an approach I pursue in the following – the distinction amounts to the availability of conditions of well-formedness on different levels of representation.

A specific writing system is never an entity of its own existence but something that is related to a specific language system and that in this sense is dependent on it. If a new writing system is to be devised, it is because there is a language that does not have one yet (or it has one, but for some reason people are not satisfied with it). Therefore, a writing system theory has to treat information of the specific language system that the writing system under consideration is related to as given information. In principle, all information of the language system is available for the mechanisms of the writing system. This does not mean that the writing system actually has to make use of all this information, of course. But there is no limit to the availability of such information.

Obviously, then, a theory of language systems is in order if a writing system theory is to be construed. What kind of information it is that is made available by the language system can only be specified with regard to a theory of language systems. Such a theory can be conceived in different ways, depending on the particular basic assumptions. In a realist reinterpretation, taking language systems as abstract objects, the gist of a hundred years of devising synchronic models of language systems in structural and generative tradition may be a model like the following:

![Figure 1: Structure of the language system (cf. Neef 2005: 5)](image-url)
Under this conception, a language system consists of two parts, namely grammar and lexicon. While grammar captures the regular aspects of the language system, the lexicon comprises everything that is irregular. In order to determine the boundary between the regular and the irregular, a theory is needed: a theory of a language system gives a reconstruction of such a system by formulating rules (or some equivalent devices) to relate data of the language system to each other. Data that by the specific theory are captured in the grammar are regular from the viewpoint of that theory (and thereby they are not only described but also explained, despite generative claims to the opposite), while data that do not conform to the rules of grammar are irregular. A characteristic of language systems is that the distinction between regular and irregular is not identical with the distinction between right and wrong (a feature, by the way, that makes argumentation in linguistics both challenging and fascinating). Among the regular forms, some are wrong, while among the irregular forms, some are correct. The regular past tense of the English verb lexeme to go is goed, which is wrong, while the irregular form went is correct. In other cases, the distinction between regular and irregular may be more disputed among theoretical linguists. From the viewpoint of the theory, the irregular form went is part of the lexicon. In this sense, the lexicon outplays grammar.

Typical examples of irregularity are simple linguistic signs in the sense of Saussure (1916). As presumably no theory can give an explanation why the concept for ‘tree’ is named as the lexeme tree in English, this arbitrary relation between form and meaning of the English lexeme tree is supposed to be part of the lexicon of the English language system. All simple lexemes are irregular in this sense and therefore part of the lexicon, while many complex lexemes show some degree of regularity and may therefore be captured – at least partly – in the grammar of the language. This conception of the lexicon as the hoard of irregularity follows Bloomfield (cf. also Neef, Vater 2006):

The form-classes of English words are largely arbitrary: there is nothing to tell us that man, boy, lad, son, father are male nouns, that run, bother are verbs, that sad, red, green are adjectives, and so on. In particular, of course, the form-class of every morpheme is arbitrarily determined. A complete description of language will list every form whose function is not determined either by structure or by a marker; it will include, accordingly, a lexicon, or list of morphemes, which indicates the form-class of each morpheme, as well as lists of all complex forms whose function is in any way irregular. (Bloomfield 1933: 269)

In Figure 1, the module grammar consists of four parts. The core capacity of a language system is to establish relations between forms and meanings. Within this general conception, the sub-component phonology is concerned with paradigmatic and syntagmatic properties of those elements that make it possible to distinguish meanings, semantics deals with relevant aspects of meaning itself, while both morphology and syntax treat formal properties of units of the language system that bear meaning. These units are words in the case of morphology and phrases (including the sentence as the ‘finite-verb phrase’) in the case of syntax. Several different theories are conceivable within this broad framework, details of which are of less interest in the context of the development of a writing system theory.

What is crucial, however, is that all elements of the language system are given information in the context of a writing system; the other components of the writing system can freely make use of it. In particular, the set of phonological segments is given information from the viewpoint of the writing system. The same holds for morphological information, the relevance of which is evident looking at the writing system of French: the words spelled <aimer>, <aimez>, and <aimée> all have the same phonological representation but constitute different (grammatical) words with different inflectional features; the spellings given refer to both phonological and morphological properties. The writing system of French functions in this way synchronically, independent of questions of its development.

For the linguist, this means that the language system has to be analyzed first before the writing system can be approached. This is what Bloomfield (1933: 21) had in mind when he put down his frequently cited statement: “in order to study writing, we must know something about language, but the reverse is not true.” The first module of a writing system is thus a language system. If a notational system is devised for something other than a language system, then it is not a writing system. It is irrelevant whether or not this language system is currently in use, as long as the writing system was originally devised for this language
system. Archeologists discovering artifacts with unknown visual patterns come to hypothesize that these patterns are signs of written language by assuming that they relate to some language system. Linguists may then use these data to try to reconstruct the language behind the writing. This way of proceeding, however, does not imply that the language depends on its written form, but the other way round.

3 The status of the script in a writing system theory

In this section, I discuss the status of a script in a writing system theory. I start with the concrete imprints on paper that are visually discernible and then proceed to the more abstract notion of script. Among the set of units of a script, I distinguish four relevant subsets. Not all of them are relevant for the writing system proper. A general outline of what may or may not be relevant is given in this section, while the next section is restricted to the notion of letter.

Arguing that a writing system is an abstract object may be counterintuitive given that what we see from written language on paper or on a computer screen are concrete objects with objectively discernible visual properties. What we perceive in this way, however, is not the writing system; we see results from the use of a writing system. From such results, i.e. empirical data, we may come to reconstruct the system that lies behind this use. This is what realist linguistics is concerned with. The linguistic sub discipline dealing with the scientific study of all aspects of written language may be termed ‘grapholinguistics’, a term suggested by Neef and Weingarten (2012ff.) as an equivalent to the German term ‘Schriftlinguistik’.

The elements used to graphically represent the units of a writing system may be called ‘glyphs’, following common parlance in typography (e.g. Lai, Yeung, Pong 1996: 308). Glyphs come in diverse shapes:

(1) Different glyphs

A A A A A A A A A A A A A A A A A

Obviously, all these single elements are visually different from each other. On a more abstract level, however, they may all be interpreted as representing one and the same unit. This is the level of characters; all glyphs in (1) may be regarded as representing the character |A| (with characters being enclosed in vertical strokes). Glyphs belong to fonts, which are studied in typography as a sub discipline of grapholinguistics. Characters belong to scripts, which are studied in graphetics (Althaus 1980, Coulmas 1999, 177-178). Although located on a more abstract level than fonts, scripts are sets of signs with visual properties or, rather, with a specific spatial organization (Harris 2009: 5). Characters of a script may show formal resemblance, be it for aesthetic reasons, or for reasons of original production depending on the writing materials used in the phase of its development, but they should not be too similar in order to make them discriminable (cf. e.g. Wiebelt 2004).

Characters can be used for different functions. In Latin, the Roman script character |l| was used for a letter as well as for a cipher (to denote the number one). It is important to distinguish such different uses. Based on the criterion of correspondence, I distinguish four different functions characters can be used for:

(2) Functional classes of the use of characters

a. letters correspond to phonological units
b. ciphers correspond to numbers
c. logograms correspond to words or morphemes (other than numbers)
d. punctuation marks correspond to (nothing linguistically)

Bredel (2011: 7-11) gives a similar typology but defines her categories with a heterogeneous set of criteria. Extensionally, her classes slightly differ from the ones given in (2).
From the viewpoint of a writing system, the most important of these classes is the class of letters. Also important are punctuation marks, while the other two classes belong to the script but not to the writing system. I will have nothing to say about ciphers here and give only few remarks about logograms and punctuation marks. I start with letters and try to delimit this notion from the notion of character.

As shown with the example of Latin, one and the same character may be used for elements of different functional classes. Within a functional class, each character can only be used once. On the other hand, an element of a functional class can make use of more than one character. This holds particularly for letters. For example, in the Roman script as used for the English writing system, the characters |A| and |a| belong to the same letter. Usually, the small character is used by convention to represent the unit letter, to be notated in angle brackets as, in this case, <a>. Altogether, the Roman script has 26 letters but uses 52 characters to represent them, with each letter having both a small and a capital character. Other scripts work differently in this respect: in Hangeul, the script used for the Korean language (cf. Lee 2009), each letter has only one character, while in the Arabic one, it has up to four. Writing systems based on scripts that supply more than one character per letter may use this variation systematically for different purposes, but this aspect is not a matter of the script itself.

Usually, letters are taken to be relevant for characterizing the general type of script under consideration. The Roman script, for example, is traditionally classified as an alphabetic script because its letters relate to phonemes (or, more general, to phonological segments). In consonantal scripts (or abjads; Daniels 2009: 29) like Hebrew, the letters relate to consonant phonemes only, in syllabic scripts like Cherokee, they relate to phonological syllables. The Cherokee script, however, is based on the Roman script (cf. Coulmas 2003, 3132). Thus, the Roman script is not inherently alphabetic but only in the context of a specific writing system. The distinction between alphabets, adjads, and syllabaries is located on the level of writing systems and not on the level of scripts.

The writing system of English currently uses the Roman script (to be given as ‘English-Roman’, following a suggestion of Weingarten 2011). The relation between writing system and a specific script is not a constituting factor of this writing system. What is actually needed for the English writing system to work is a script that supplies 52 characters, to be grouped into 26 pairs of characters as letters (to give the alphabet of the English writing system). Which script is actually chosen for the writing system of English is a matter of orthographic convention, something that lies outside the scope of a writing system theory. The situation with the German writing system is slightly different. This system uses the German script as a modification of the Roman script (‘German-Roman”). Four letters are added to the original set, namely the umlaut letters <ä>, <ö>, and <ü> plus <ß> (a letter that, incidentally, does not have a capital character). What is crucial to the current point is that the characters of the umlaut letters are the characters of the vowel letters <a>, <o>, and <u>, respectively, added by a trema. This relation in shape can be regarded as a necessary condition on the script because certain orthographic regularities can be analyzed as making recourse to this feature (regularities pertaining to the fields of stem constancy on the one hand and the absence of doubled umlaut letters on the other).

If a writing system was in need of a script, this script could be invented from scratch or it could be borrowed from a writing system of a different language. The latter case is the one that occurred much more often in history than the former one. In principle, a thorough analysis of the language system could lead to arguments in favor of a specific type of script, while in practice, arguments pertaining to the specific historical context of the society tend to determine the choice of the script to borrow (cf. Unseth 2008). In such a situation, it may turn out that the script chosen does not sufficiently match the purposes of the writing system under development. Daniels (2006) describes different reactions that are possible under such circumstances, restricting his view to the set of letters:

Scripts are often borrowed or adapted for writing new languages, and the borrowing language usually includes sounds not found in the source language. Mechanisms for accommodating new sounds or phonotactics have not been studied as a group before, and a wide variety of cases is considered here. The techniques are found to fall into a limited number of categories: inventory reduction, inventory expansion, character combination, character alteration, character borrowing, systematic additions to characters, diacritics, and character simplification. (Daniels 2006: 7)
So far, I have restricted the discussion to scripts used for phonographic writing systems. Their basic units are letters. Logographic writing systems differ from phonographic writing systems in that their basic units are logograms, according to the terminology given in (2), i.e. functional classes that correspond to words or morphemes. Phonographic scripts typically also have elements with logographic properties. Examples in use of the English writing system are ampersand \(|&|\), section sign \(|§|\), and degree \(|°|\). The set of such symbols is difficult to determine because of the many respective examples in technical terminology. It might be worthwhile to restrict the analysis of writing systems to everyday language, but the division line between standard vocabulary and technical terminology is notoriously difficult to draw. Logographic scripts, on the other hand, may also have elements that relate to phonological units, a feature much discussed with regard to the Chinese writing system (e.g. Boltz 1996, Mair 1996; see also Unger 2011).

There are also characters that do not relate to units of language proper. A first set of written signs that may come to mind in this context are diacritics, but these are more adequately regarded as parts of characters, not as characters in their own right. True examples for the type of characters indicated are punctuation marks like apostrophe \(|'|\), brackets \(|( )|\), and question mark \(|?|\) as well as spaces. Writing systems using scripts with punctuation marks have several options for functionalizing such elements (cf. Nunberg 1990, Bredel 2008).

### 4 The second module: graphematics

The component of the writing system that captures the relation between letters and phonological units of the language system is what I call graphematics (cf. Neef 2005). As is widespread for linguistic terms, there are several different, though related, concepts connected with the designation ‘graphematics’, and there are alternative designations for at least some of these concepts (cf. Neef 2012). In German linguistics, in particular, the term ‘graphematics’ is currently mainly in use for a concept that resembles what I call ‘writing system’ (e.g. Eisenberg 2013: 287, Dürscheid 2012). Approaches that pursue such a broad concept of graphematics principally neglect the distinction between two different modules within a writing system, modules I call ‘graphematics’ and ‘systematic orthography’, respectively (cf. paragraph 6 below).

Different theories are conceivable to model the graphematics module of a writing system. A specific approach is the **Recoding Model of Graphematics**, originally devised for the German writing system, but transferable to other systems as well (Neef 2005). The following explanations are framed in this theory in which graphematics is conceived of as a rule system. The role of graphematics in this sense is to relate phonological representations to any given sequence of letters. This is what is called ‘recoding’ in this approach. For each single letter, graphematics has a rule stating which phonological unit(s) of the language system it corresponds to. Thus, I focus on the new information, i.e. the set of letters, and see how it relates to the given information, i.e. phonological units of the language system. Letters are, thus, defined by their individual correspondence to phonological units of the language system. Features of scripts that are relevant for the writing system are in this conception captured in the concept of letter.

The tradition in grapholinguistics that gives precedence to the perspective ‘from letter to sound’ is first and foremost connected with work by Richard Venezky (1970, 1999). This approach has found rather few followers, though. A majority of theories in grapholinguistics takes the opposite direction as basic; such theories give correspondence rules for phonemes and more or less neglect or downplay the relevance of the other direction (e.g. Eisenberg 2013: 291, Nunn 1998, Sproat 2000). Only few studies like Nerius (2007) for German and Carney (1994) for English pay equal attention to both directions. Besides a number of theoretical problems connected with the approach ‘from sound to letter’ (cf. Neef 2004: 138-143, 2010: 13-16), it is unexpected in such a model that for some phonological units like stress – at least in many writing systems like English and German – there is no corresponding written unit. In the end, it is evident that a theory of writing systems has to model regularities in both directions. The specific approach defended here takes the perspective ‘from letter to sound’ to be basic (for reasons to be given later) but regards the other perspective as important as well. Each of the two perspectives is modeled in a module of its own, namely graphematics and systematic orthography, respectively.
Graphematics, thus, deals with the relation of letters to phonological units. Such relations may be simple in that a certain letter is to be recoded exclusively as a specific phonological unit. A relation is complex if there is more than one phonological element that regularly corresponds to a letter. In this case, the actual correspondence may be determined by the graphematic context or by phonological regularities (regarded as given information), but it may also be left undetermined (cf. Neef, Balestra 2011 for lists of correspondence rules for the writing systems of German and Italian). In alphabetic writing systems, the element that corresponds with a letter may be a single phonological segment, but also combinations of such segments like diphthongs (e.g. <cry> → [kraɪ], to be read as ‘the spelling <cry> is recoded as the phonological representation [kraɪ]’), affricates (e.g. <jam> → [dʒæm]), and combinations without a specific phonological status (e.g. <box> → [bɔks]) are possible. A letter may also allow the correspondence to a specific phonological unit in some contexts but to nothing in others. An example of the English writing system is the letter <h> that corresponds to [h] in some cases like <hair> but to nothing in others like <verandah>. Moreover, specific combinations of letters may behave as units with respect to their correspondences. An example is English <th> that in most instances corresponds to [ð] or [ɵ]. Such combinations of letters are termed ‘fixed letter combinations’ in the Recoding Model; they add to the list of graphematic correspondence rules of a writing system.

Finally, the graphematics component of a writing system may be further equipped with graphematic constraints that exceed individual correspondence rules of letters. A constraint widespread in writing systems using the Roman script allows a zero correspondence for any letter that follows an instance of the same letter. A good example is the English spelling <shimmer> (for more details concerning this particular constraint, cf. Neef 2005: 78-81, Neef 2012: 221 for German, Neef, Balestra 2011: 127-128 for Italian).

The above considerations lead to the following figure of a writing system with its obligatory components ‘language system’ and ‘graphematics’. This kind of writing system lacks orthography and is, thus, of the graphematic type. In the next section, I will consider whether such a writing system is ready to work. In the diagram, the colors are meant to indicate that the language system has an existence independent of the writing system while graphematics does not.

![Diagram of writing system with language system and graphematics](image)

**Figure 2: Structure of the writing system (graphematic type)**

## 5 A case study: English-IPA as a writing system

The attempt to develop a theory of phonographic writing systems leads to the assumption that for doing so, two modules are needed: language system and graphematics. This latter component by necessity includes a set of letters with one correspondence rule for each letter. In this section, I want to consider whether there are writing systems that can be sufficiently analyzed based on this minimal set up. In particular, I want to argue that the writing system English-IPA is such a graphematic writing system.

---

2 This statement depends on the specific phonological theory chosen. It may well be the case that there is a phonological theory that regards [ks] in English as a combination with a specific phonological status.
The International Phonetic Alphabet (IPA), though originally devised as a phonetic transcription system, is more aptly characterized as an alphabetic script (cf. Sampson 1985: 33, Coulmas 2003: 28-33, Harris 2009: 54). In particular, IPA does not constitute a mechanical procedure to reduce the continuous sound stream of spoken language to discrete symbols, as one would expect from a system for phonetic transcription (either broad or narrow). With its use of discrete symbols, IPA is closer to phonology than to phonetics. Therefore, IPA may be regarded as a means to give phonological transcriptions or, more aptly, phonological representations (cf. Vennemann, Jacobs 1982: 36).

This evaluation also follows from an adequate interpretation of the basic premise of IPA, which reads as follows:

When two sounds occurring in a given language are employed for distinguishing one word from another, they should wherever possible be represented by two distinct symbols without diacritics. (IPA 1999: 159)

This ‘phonemic principle’ (IPA 1999: 29), which is obviously only relevant for ‘broad’ or ‘phonemic transcriptions’, implicitly gives a definition of the term ‘phoneme’ (cf. IPA 1999: 27). Furthermore, it states that in an IPA-based writing system, there is a unique correspondence relation from phoneme to letter, given that the term ‘symbol’ in the citation can be equated with the term ‘letter’ as defined above. The qualification that ‘wherever possible’ simple symbols should be used for representing phonemes actually requires dispensing with the colon to indicate vowel length in languages like English and German where short vowels and long vowels have the status of distinct phonemes (at least according to some phonological theories, cf. Hall 2011: 68-69 for German; Meyer et al. 2005: 90 for English). In practice, however, most users of IPA mark length by a discrete colon (but cf. IPA 1999: 30 where representing the different vowel phonemes in the English words bid and bead with the simple symbols [i] and [i] – actually given in slant brackets – is explicitly allowed). From the original formulation of the first principle from 1888 (“There should be a separate letter for each distinctive sound”; cf. IPA 2010: 357) it follows that each letter has to be used for representing one phoneme only, presumably also implying that each letter has to have exactly one character only to represent it. Together, these principles lead to a one-to-one relation between letters (or characters) and phonological segments in both directions.

Following from this interpretation, I regard IPA as a model kit that can be used to create writing systems for each single language. The IPA script supplies a large number of characters plus additional symbols of different status. In order to deploy IPA for a specific language, only a subset of these characters is needed. Therefore, the specific IPA-based writing system for English should more aptly be given as ‘English-IPAEng’, indicating that the writing system of English uses a modified version of the IPA-script, more precisely a specific reduced version or subset, as is the case for each IPA-based writing system. In an IPA-based writing system, the set of letters matches the set of phonological segments, all correspondence rules are simple, and neither fixed letter combinations nor graphematic constraints exist. Nevertheless, such a writing system allows writing any text that is conceivable based on the particular language system. It is, thus, a perfectly viable writing system. It is most shallow (in terms of the concept of orthographic depth, cf. Katz, Feldman 1983) and absolutely transparent (in terms of Neef, Balestra 2011).

Such a purely graphematic system may or may not be regarded as an ideal type of a writing system (cf. Coulmas 2003: 26-33, Venezyk 2004). Natural writing systems, however, evidently work differently. The greater complexity of natural writing systems does not only follow from a more complex component of graphematics, but from an additional module, to be discussed in the following section.

6 An optional (though widespread) module: systematic orthography

With the two modules established, the writing system is ready for use: through the interplay of graphematics and regularities of the language system, any sequence of letters pertaining to a specific (phonographic) writing system can be related to at least one phonological representation. This representation may be either
grammatical as is the case for the English spelling <try> that regularly corresponds to [traɪ] or ungrammatical as would hold for the spelling <rty> in the English writing system, presumably to be recoded as *[traɪ]. At the same time, it is clear how a specific word of the language system has to be spelled, namely in any way that allows the recoding of the phonological representation of the word in question. In other words: if the graphematics of the English writing system provides the recoding of <try> as [traɪ], then <try> would be a graphematically licensed spelling of each word with the phonological representation [traɪ]. In such a way, graphematics defines what a possible spelling of a word is. In IPA-based writing systems, there is always exactly one possible spelling for a word; in natural writing systems, there may be more than one, actually considerably more than one. The set of possible spellings for a word with a specific phonological representation is what I call the ‘graphematic solution space’. The following list gives an excerpt of the graphematic solution space of the phonological representation [raɪt] as a subject of the writing system of English:

(3) Graphematic solution space of the phonological form [raɪt] in English (excerpt)
   a. right
   b. rite
   c. write
   d. wright
   e. wrightt
   f. ryte
   g. rrryte

The writing system English-Roman, as opposed to English-IPA_english, has a rich graphematic system with complex correspondence rules, fixed letter combinations, and graphematic constraints, as is typical of natural writing systems. This leads to large graphematic solution spaces for specific phonological representations, which, in turn, allow massive variation in spelling. Such variation is not necessarily a pointless complication. It allows, for example, different spellings of homophonous words like right and rite, and it allows the constant spelling of morphemes that show phonological variation like the plural morpheme –s (based on an underdetermined correspondence rule of the type ‘<s> → [s] v [z]’).

Now, in English the word right is not allowed to be spelled as <rite> although this spelling is part of the graphematic solution space of the relevant phonological representation [raɪt]. Obviously, graphematics is not enough to explain the actual spellings of words. Here, a further module of the writing system comes into play, a component that prescribes how to write correctly within the limits of the graphematic solution space. This module particularly aims at constant spellings of words. Such a component may for obvious reasons be called ‘orthography’. How could orthography work in the present conception? There could be a norm-giving authority that in each single case decides which of the elements of the graphematic solution space should be the conventionally correct spelling of a word in question. Such an authority may now and then even declare a spelling as correct that is not even element of the graphematic solution space. In such a case, a language would have word spellings that are irregular. Irregularities of this kind are in fact typical of natural writing systems.

The factual instantiation of this norm-giving authority is rather the whole writing community in the sense of an ‘invisible hand’ (cf. Keller 2014) than a specific person or group of persons. In any case, it is probable that this authority makes its decisions not in an absolutely random way but that it follows some precepts or guidelines in selecting the specific spellings of a word from its graphematic solution space. It is the task of the linguist to detect these precepts and to reconstruct them as a system. Taken together, these guidelines form an optional component of the writing system theory which I call ‘systematic orthography’. This component is to be distinguished from ‘conventional orthography’ that is captured in explicit norms like the ‘Amtliche Regelung der deutschen Rechtschreibung’ with respect to German.

---

3 I refrain from using the term ‘principle’ in this context because this term is tied to a specific tradition of modeling orthography, characterized by the use of terms like ‘phonological principle’ and ‘morphological principle’, to mention but a few. Problems of this theoretical approach are discussed in Neef (2013).
Obviously, spellings such as <rrrryte> (cf. (3)) are almost impossible in English, but they may exist in contexts when language is played with like in comics, for example (cf. Forster, Borgwaldt, Neef 2012). Specifically, written words must not begin with more than one identical consonant letter unless the word is an expressive spelling or an interjection. This regularity can be captured in a constraint that applies to the graphematic solution space and that is effective for specific levels of the vocabulary only. ‘Vocabulary’ is a different concept from ‘lexicon’ as sketched above in that it is simply the set of actual words or lexemes of a language, irrespective of the distinction between regular and irregular. In a theoretical way, the vocabulary of a language can be divided into different levels. These levels are relevant for grammar as well as for the writing system. Levels of the vocabulary mainly come into view when discussing the status of foreign words; cf. Eisenberg (2011).

The standard approach is to regard words as foreign that have an approved foreign origin. The approach chosen here is different because for a synchronic model it is more important to evaluate present properties of words than to refer to the historic origin, which in fact hardly ever is to be traced back to an ultimate beginning. Under this view, the German word Sport ‘sports’ is a native word because it is fully assimilated although it was borrowed from English in the 19th century (Kluge 2011: 870).

The following model was originally devised for German (cf. Neef 2005: 205, Balestra, Appelt, Neef 2014: 136, Appelt, Balestra, Neef 2015: 157) but may be applied to English and other languages as well:

![Figure 3: Levels of the vocabulary of German](image)

The ‘Whole Vocabulary’ consists of foreign-language words plus the German Vocabulary. Foreign-language words are words that have grammatical properties that follow the grammar of the source language but violate aspects of the German grammar. Thriller, for example, is such a word in German as it begins with a sound that is not a phoneme of German. The ‘German Vocabulary’ consists of interjections and other marginal word classes plus the Grammatical Vocabulary. Interjections are German in that they have German phonemes only, but they do not fall in the scope of grammar in that they do not have to obey relevant rules of German. The interjection pst, for ‘shut up’ does not have a vowel, which is impossible otherwise for words of German.

The ‘Grammatical Vocabulary’ consists of indigenous proper names plus the Core Vocabulary. Words of the Grammatical Vocabulary follow the rules of grammar, but words in the center of the Grammatical Vocabulary have to adhere to more rules than words at outer shells. For example, in indigenous proper names (including complex words based on such names), the vocalic r [ʁ] as a syllable peak is allowed to immediately follow an...

---

4 There is a related, though distinct concept of ‘levels of the lexicon’ in a multitude of linguistic theories, most notably in Lexical Phonology (cf. Kiparsky 1982).
unstressed full vowel. Therefore, an inhabitant of the city of Jena may be called Jenaer [ˈje.nɑ.ɐ]. This phonological constellation is not allowed in the Core Vocabulary. Therefore, the comparative form of the adjective rosa ‘pink’ must not be *rosaer [ˈro.za.ɐ] although this form follows the normal rules of formation of comparative forms (cf. Neef 1999). The ‘Core Vocabulary’ consists of assimilated foreign words plus the Native Vocabulary. The phoneme [ʒ] is allowed in assimilated foreign words like Garage [ˈɡa.ɾɑ.ʒə] or Dschungel ‘jungle’ [ˈdʒʊŋ̣əl] – and, of course, in words of all outer levels of the vocabulary – but not in native words.

The German writing system refers to the same classes of the vocabulary: it has nothing to say about foreign language-words; all words of the German Vocabulary are subject to graphemics, but only words of the Grammatical Vocabulary are subject to systematic orthography. Native words follow more orthographic constraints than assimilated foreign words and both types follow more constraints than indigenous proper names. For example, foreign language words may contain letters that do not belong to the German alphabet like Façon ‘style’ with the French letter <ç> (e.g. in Duden 2001: 511); only interjections are allowed to start with more than one identical consonant letter (mmmmh); foreign words may contain the letter sequence <th> to correspond with [t] like Theater ‘theatre’ [te.ˈa.te], native words may not.⁵ As a result, the writing system gives more reliable specifications for the spelling of native words than it does for words at outer shells of the vocabulary. A proper name with the phonological representation [hut] may be spelled with final <th> or simple <t> as <Huth> or <Hut>; the homophonous native word Hut ‘hat’ can only be spelled with a simple <t>.

Tentatively assuming that English systematic orthography works essentially in the same way, the status of words like Lloyd and Fforde can be considered. If classified as foreign-language words, the spelling of these words is approved by the English writing system in that this system does not impose any constraints on the spelling of words of this level. It is also possible and maybe even more convincing to regard the cases in question as indigenous proper names. Since words of this level of the vocabulary are not allowed to start with two identical consonant letters, the spellings of both these words are wrong from the viewpoint of the writing system. Still, these spellings are correct from the viewpoint of conventional orthography. A theory like the one sketched here, thus, supplies a measure to mark spellings as irregular. If the latter evaluation is convincing, spelling reformers may have the impulse to regularize the spellings under discussion to <Loyd> and <Forde>, respectively (if they ever dare to take proper names into account).

To sum up, systematic orthography may be construed as a system of constraints applying to the graphematic solution space. These constraints are sensible for different levels of the vocabulary. The relevant levels are already given in the language system module. In general, systematic orthography is an optional component of the writing system, which would be solidly designed without this module (as is the case for IPA-based writing systems). Adding systematic orthography to a writing system turns this system from the graphematic type into the orthographic type. In a sense, systematic orthography is dependent from graphemics (via the graphematic solution space). This dependency is captured in the following figure:

---

⁵ This statement is true since the implementation of the German standard orthography in 1901.
7 Conclusion: on the relation of spoken to written language

If the term ‘writing system’ is defined as a notational system for a language system – which is appropriate from the viewpoint of theoretical linguistics – it follows that the language system has to be conceived as being part of the writing system, with both language system and writing system being construed as abstract objects in the sense of Linguistic Realism. What essentially constitutes a (phonographic) writing system is graphematics as the component that comprises letter rules, i.e. correspondence rules that relate the units of the writing system to phonological units of the language system. Purely graphematic writing systems are sometimes declared to be the ideal of a writing system, though in reality they hardly exist. The application of the script IPA to a specific language system can be interpreted as such a type of writing system, a first order (or graphematic) system, so to say.

The normal case of a writing system is a second order (or orthographic) writing system, a system that besides graphematics has a further module called systematic orthography. In the theoretical framework of the Modular Theory of Writing Systems, this module is construed as a selection procedure applying to the graphematic solution space of specific phonological representations. Systematic orthography consists of constraints that are sensitive to specific levels of the vocabulary of a language. With an appropriate methodology, linguists are able to reconstruct the regularities that are fundamental to conventional orthography in the form of a theoretical model of an orthographic writing system. Such theorizing allows distinguishing among the conventionally correct spellings between the regular and the irregular ones. This may further serve as a basis for didactic models and reforms of orthography.

The above reflections on the construction of a writing system theory allow an answer to the much debated question how spoken language and written language are related to each other. A popular conception of this aspect is that written language is derived from spoken language, or spelling is derived from speaking. This view can also be found in more sophisticated work like Daniels (2009: 36) who states: “Since writing represents language, it must represent the sounds of speech.” If taken literally, this argument is not conclusive because it implies that language is speech. However, trying to derive spelling from speech is a pointless venture. Spoken language consists of continuous sound streams that hardly give a basis for the derivation of spellings with discrete units. Moreover, spoken language shows massive variation, written language does not.

If such a speech-based approach is considered seriously in grapholinguistics, the input of the derivation is not taken to be the result of some phonetic measuring but a (broad) phonetic transcription. Such a transcription, however, is not a representation of spoken language in an empirical sense of phonetics but, as argued above, merely a different kind of writing system. In the end, so-called speech-based approaches attempt to translate representations of a graphematic writing system into representations of an orthographic writing system, whatever kind of scientific procedure that is.

On a more systematic level, the current question can be reformulated as concerning the relation between language system and writing system. Eisenberg (2013: 288) stipulates that a writing system is part of grammar. In the same vein, from its fourth edition (1984) onwards the popular German Duden-grammar has a chapter on ‘the letter’. A different conception of the relation in question is given in Neef (2005: 5) where I argue that writing system and language system stand side by side but are basically independent of each other. Now I see it differently: the language system is a module of the writing system, as argued for above.

From this theoretical conception it follows that written representations do not relate to the phonetics of a language, as phonetics is situated outside the language system and pertains to the realm of language use. Instead, written representations are related to phonological representations. The relation between these two kinds of representations, however, is not derivational but rather indirect: a regular written representation of a specific word is element of the graphematic solution space of the phonological representation of that word. To put it differently: a word has to be spelled in a way that allows the regular recoding of its phonological representation. If the graphematic component of a writing system supplies more than one

---

6 In his terms: ‘graphematics’, a different term from my one with the same designation; cf. section 4.
such spelling option, as is the normal case in natural writing systems, a set of systematic-orthographic constraints may reduce the number of options. This reduction does not necessarily leave one fixed spelling of the word but usually a larger set. Conventional orthography in the end decides which of these options the correct one is. This means that conventional orthography cannot be fully reconstructed as a theoretical system, but only partially.

Acknowledgments: This work grew out of the project ‘Systematic Orthography of German’, supported by the Deutsche Forschungsgemeinschaft (DFG), Grant No. NE 857/2. I like to thank my colleagues in the project, Annalen Appelt and Miriam Balestra, for stimulating discussions as well as Monika Budde and a number of anonymous reviewers who helped me to clarify my thoughts.

References

Writing systems as modular objects: proposals for theory design in grapholinguistics


