

PETER EISENBERG

Writing system and morphology. Some orthographic regularities of German

Abstract

There are detailed orthographic rules for most alphabetic writings, which tell us in nearly all cases how to write correctly. Yet it is not very well known, to what extent these rules correspond or even coincide with the linguistic rules underlying the writing system of the language in question.

The approach outlined here presupposes that the graphemic part of a grammar is to be considered as an integral part of the overall language system, and that it has to be investigated by the same linguistic methods as the other subsystems. Support is given for this thesis by an examination of the relation between morphology and graphemics in German. It is argued that there is a close mutual dependency between the morphology and the writing system.

1. Remarks on the linguistic level of alphabetic writings

There is some agreement among linguists working on problems of writing systems, that these systems have to be considered as proper subsystems of language systems and that it should be desirable, therefore, and possible, to treat the graphemic part of a grammar as one of its components in the usual sense, i.e., to fully integrate it into an overall grammar of the language under description. There is much less agreement about the linguistic level on which the graphemic component should be based, and this paper is not intended to make any substantial contribution to this problem. Nevertheless, I would like to make some remarks about my own view of the present situation. These remarks serve the purpose of clarifying what is intended by the proposals made in Sections 2 and 3, and thereby of avoiding counter-arguments which do not meet the point I want to make.

With respect to the linguistic level of alphabetic systems, the most debated question is whether, and to what extent, these systems are phonologically

or phonetically based, and it has been shown several times that much confusion exists concerning this problem.¹ In my opinion, most of this confusion originates from the simple, if not trivial, fact that there is much discussion about the principles of orthography without sufficient agreement about the concepts of phonology and phonetics used in these discussions. But even if it is clear what should be understood with respect to phonology vs. phonetics, people sometimes tend to generalize certain observations, regarding them as 'proof' of the fact that the whole system should be taken to have this or that property as its basic property. So we read statements like „daß die deutsche Orthografie sich direkt auf die Phoneme/ Grafeme bezieht und nicht auf phonetische Varianten, ist am velaren Nasal [ŋ] nachzuweisen" (Müller 1978: 28), where the velar nasal is taken as evidence for the 'phonemic principle' in German orthography. I think William Haas (1975 and in this volume) has argued quite convincingly that it is not justifiable to place existing writing systems on one linguistic level in such a way that it could not have access to information from other levels. On the other hand it does not, according to Haas, alter the level of a script if it takes into account information from any other level. A syllabic script remains syllabic even if it depends in part on the phonetic level and vice versa. Similar conclusions can be drawn from what still appears to be the most conclusive approach to graphemics, especially its integration into an overall grammar, namely the one outlined in Bierwisch 1972.

Bierwisch presupposes a generative phonology of the Chomsky/Halle 1968 type. The relation between the sound level and the graphemic level is realized by so-called grapheme-phoneme-correspondence rules (GPK rules). These are context-sensitive rewrite rules which formally resemble very much the usual phonological rules. They differ essentially from phonological ones in not containing units from any sound level on their right side, but units from the graphemic level. Each GPK rule allows for the rewriting of a sequence of sound units by a sequence of graphemic surface units, i.e. letters. Bierwisch is able to show that in German not all GPK rules can operate on the deep phonological level. Some of these rules can only operate after certain phonological rules have been applied, but as it is not possible to give any generalized criteria for the graphemic relevance of phonological rules, it follows that it might only be possible to follow certain GPK rules after all phonological rules have first been applied, i.e. on the phonetic level.

Despite the fact that Bierwisch uses – and given his approach has to use – a fully elaborated generative phonology, it is not at all clear whether we should speak of our writing system as 'phonological'. The reason for this is what has just been said about the level of application of GPK rules. There is

no level on which they operate, and *a fortiori* it is not the phonological level.

Even though the following arguments will show that I have largely adopted Bierwisch's way of thinking about the writing system, I want to stress, nevertheless, the term 'morphological', and do not make any attempt to integrate what I have to say into a generative phonology. There are essentially two reasons for this. First there are facts to be handled which cannot be handled, as far as I can see, in a reasonable way within generative phonology. As an example, let me mention the relevance of the morphological distinction productive/non-productive for graphemics. Secondly, there seems to be a tendency, at least in the lexicalist quarters of the generative camp, to establish a morphological component within the grammar, and to pay more attention to the morphological and, especially, to word formation rules in their own right, than was the case in earlier periods. This would mean that approaches of the Chomsky/Halle type, which did not leave much room for morphology between syntax and phonology, should be abandoned (see for instance Aronoff 1976 and Booij 1977). The design of the morphological part of the grammar then follows the still more general tendency not to separate the different subsystems completely, but to provide the possibility of mutual interaction. It seems to be the case, furthermore, that practically all components of the grammar do heavily rely on information from the surface level (with respect to morphology cf. Motsch 1977 and especially Plank 1980).

The graphemic subsystem of German gives us some hints pointing in the same direction. So it can easily be shown that for the purpose of grasping the regularities of graphemics one should use not only phonemic and morphosyntactic, but also graphemic information itself which, in turn, serves morphological purposes (see Section 2). This certainly has to be taken into account if one reflects the possibilities to integrate graphemics into the grammar. It is not our goal to do this here. By discussing some examples, we only want to show what kind of influence of morphology on writing one has to be aware of.

2. *The principle of greatest similarity*

It is one of the most obvious properties of our writing system that the graphemic representations of different units from a paradigm are made as similar as possible to each other. This fact is normally accounted for and explained by stating that graphemic representations are in many cases not derived from phonetic, but from some deeper structures. Therefore, certain phonetic differences between units of the same paradigm might be neglected within the written forms of these units. Since the mechanism in question does not

provide in all cases *identical*, but very often only *similar*, graphemic representations, we propose to call it the *principle of greatest similarity* and not the principle of identity, as one could first be inclined to do. Some relevant examples that come immediately to mind are given below.

The principle of greatest similarity has a systematic and a historical aspect. According to a widespread opinion, it is this principle which is to some extent responsible for the growing influence of written language on the native speaker's knowledge about his language. Since, by this principle, the writing system directly reflects the morphological structure of the language, and since writing systems seem to be more conservative than sound systems, it follows that one can expect a direct interaction between the functioning of the writing system and morphological change. More specifically, one can expect that the writing system tends to prevent certain morphological changes since it keeps alive the native speaker's knowledge about those derivational relations, which could easily be lost if this knowledge had to be based on the sound system alone. Examples of this kind are available from the morphological literature. Plank, for instance, (1980: 95ff.) demonstrates that morphological reanalysis, which is one of the most important mechanisms of morphological change, would occur more often than it does if we did not have a writing system of the kind we have.

Let us now illustrate some of the effects of the principle of greatest similarity in German orthography. In German, as in English, a suffix *iv* is used to derive adjectival stems from noun stems, as, for example, in *extensiv*, *kooperativ*, *massiv*, *produktiv*. For the stem of these adjectives, one could think in principle of three different graphemic representations with respect to the suffix *iv*, which would all be in accordance with the principle of greatest similarity, namely:

- (1) a. *massiwer*–*massiw*
- b. *massifer*–*massif*
- c. *massiver*–*massiv*

(1a) would be the 'normal' phonological spelling in German. Since this is a case of final devoicing, and since this rule is generally neglected by the GPK rules (i.e. the voiced consonant is the only one which appears on the graphemic level), one could expect ⟨w⟩ for both [v] and [f].² (1b) could be motivated by the French writing. This would mean that we had taken over the French form ⟨*massif*⟩ and maintained this form in German, though it would not fit the German system. Yet it is clear that we often accept foreign spellings which do not fit our system when we accept foreign words. (1c) could be motivated by similar reasons, since in French we do have ⟨*massive*⟩ as well as ⟨*massif*⟩.

The reason why the German system in fact selects (1c) seems to be the following. We do not have a phonological rule of final devoicing which replaces /v/ by /f/, though this would be possible with respect to the phonetic values of these segments. /v/ and /f/ constitute one of the two or three pairs of consonants in German which meet the phonetic conditions of final devoicing, but which, nevertheless, are not related to each other by this rule. Therefore, the system cannot accept (1a). Similar for (1b): there is no case in German where a voiceless consonant could be read as voiced. Originally, this was also true for ⟨v⟩, which has as its usual reading a voiceless consonant such as the ⟨v⟩ in *Vater*, *viel*, *Gustav*, *verlieren*. Yet in another group of words, namely foreign words, ⟨v⟩ as a historically younger reading has also the value /v/, as in *Version*, *Vene*, *Verdikt*. This means that ⟨v⟩ is the only one of the three letters for which we have a voiced and a voiceless reading, though these readings are never related to each other by final devoicing. This seems to be the reason why the writing system accepts ⟨v⟩ for all forms of the paradigm. In doing this, another regularity must clearly be neglected, namely the one which says that in case of final devoicing, the writing system is based on a deeper phonological level, i.e. it chooses the letter which normally stands for the voiced consonant. This example illustrates that the principle of greatest similarity is even maintained in certain cases where other important rules are canceled.

From the following examples, it can be seen that our principle is indeed one of similarity, not of identity. In these cases, morphological alternations correspond to differences in spelling, but these differences are smaller in the written form than in the phonetic and phonological form. We are concerned with vowel alternations in German, which are conventionally divided into the subclasses Umlaut and Ablaut. As far as I can see, this terminology is not motivated phonetically but graphemically, since I do not know any phonetic criteria which would allow the introduction of a special class of pairs of phonetic units, consisting each of 'Laut' and 'Umlaut'.

The principle of greatest similarity is realized with respect to Umlaut by introducing new letters into the alphabet which constitute a formal similarity between units not existing on the phonological or phonetic level. For the diphthong [oi] this even means that a graphemic representation ⟨äu⟩ is introduced, which has the same phonetic counterpart as the ⟨eu⟩ used 'normally' and which is only used to establish a formal similarity with respect to ⟨au⟩. In other words, ⟨äu⟩ is only used within units which are (by inflexion or derivation) morphologically related to units with ⟨au⟩ in the corresponding position, such as in *Haus*–*Häuser*, *Auge*–*äugen*. ⟨äu⟩ is never used in any other position, therefore it never occurs in words such as *euch*, *Efeu*, *Freund*. It

should be clear that in cases of this kind, graphemic similarity goes beyond similarity on any level of sound.

There are other means of expressing similarities with respect to Ablaut, because we find them in the occurrences of vowels within the forms of so-called strong verbs and mixed verbs. Problems of economy would arise if the system used similar letters to express the paradigmatic relations between forms such as *trinke*–*trank*–*getrunken*. It would cause an inflation of graphemic ambiguities and lead to a very complex system of graphemes if we invented special letters to express the fact that [i], [a], and [u] appear in the same position within the forms of a paradigm. But the possibility for expressing similarity does exist with respect to the graphemic realization of vowel length as it is expressed by the occurrence or non-occurrence of the so-called 'Dehnungs-*h*'.

The rules about where the Dehnungs-*h* cannot appear are pretty clear and well-known. These rules are based on purely phonetic and graphemic conditions. In contrast to this, we do not have any rule which tells us, where the Dehnungs-*h* must appear. Nevertheless, it always holds that if the Dehnungs-*h* appears in one form of a verbal paradigm to mark a vowel as [+tense], it will also be used in all other forms of that paradigm if a tense vowel occurs in that position, even if we have different vowels. The same holds if the tenseness is not expressed by the Dehnungs-*h*. So we have *stehlen*–*stahl*–*gestohlen*, *fahren*–*fuhr*, *befehlen*–*befahl*, but *gebären*–*gebar*–*geboren*, *gären*–*gegoren*, *küren*–*gekoren*. I know of no exception to this principle.

The consequence of this regularity is especially impressive in the case of [i:]. In many orthographic books we find a rule saying that ⟨e⟩ has a similar function with respect to ⟨i⟩ as has the Dehnungs-*h* with respect to the other vowels, i.e. [i:] is graphemically realized as ⟨ie⟩ instead of ⟨ih⟩. This does not hold for those cases which we are considering. The Dehnungs-*h* is *always* found in forms with an [i:] (i.e. a tense vowel) if it appears in other forms of the paradigm too, and of course it then never appears only in forms with [i:]. As a consequence, we get three graphemic representations of [i:]. We write *Igel* and *viel* as we write *Abend* and *kahl*: [i:] is represented as ⟨i⟩ and as ⟨ie⟩ just as [a:] is represented as ⟨a⟩ and ⟨ah⟩. We then have, of course, ⟨ie⟩ according to the principle of greatest similarity where [i:] is in the relation of Ablaut to another tense vowel represented without Dehnungs-*h*, such as in *verlieren*–*verlor*, *rufen*–*rief*. On the other hand, we have ⟨ieh⟩ as graphemic representation of [i:] only in those cases, where [i:] is in the relation of Ablaut to another tensed vowel which is represented graphemically by vowel plus Dehnungs-*h*, as such in *befehlen*–*befiehlt*, *empfehlen*–*empfiehl*, *stehlen*–*stiehlt*. In this respect, ⟨ie⟩ behaves like a simple letter. There are very few cases in German where [i:] is represented as ⟨ieh⟩ for other reasons than the

ones just stated, such as in *Vieh* and its derivatives. These are not just exceptions to our rule, but there are special etymological reasons for such spellings which we cannot go into here. There is a fourth graphemic representation for [i:], namely the one such as ⟨ih⟩, in, for example, *ihr*, *ihn*. It has an interesting, but very limited distribution and is not our concern here.

From this example, it can be seen how consistently paradigmatic relations are expressed by the writing system. Graphemic similarities of this kind seem to be very resistant to historical change, and, of course, do not have any direct correspondence at the sound level.

3. Morphological determination of graphemic ambiguity

Before we try to explain why certain types of graphemic ambiguities occur in our writing system, it seems to be useful to comment briefly on the concept of graphemic ambiguity itself. When we use the term 'graphemic' in this paper, we always refer to the surface level of written texts. On the surface level, we are concerned with the letters of our alphabet and the rules for their combination. It is not our aim to contribute to or make use of a 'graphemic theory', which classifies letters and sequences of letters, the 'graphs', into classes called 'graphemes'. A theory of this kind would have to consider graphemic structures not as surface structures, but as structures on some deeper level of description. I do not want to discuss any of these theories here, but only want to make clear that we are always talking about the surface of scripts, considered as consisting of sequences of letters.

It seems reasonable, then, to base the concept of graphemic ambiguity entirely on the surface level, i.e. to relate the surface of written texts to the surface of spoken texts and not to any deeper phonological level. By doing this, we are following Bierwisch (1972: 75), who considers a *phonetic* representation as being *graphemically* ambiguous, if and only if it has at least two graphemic representations. According to Bierwisch (*ibid.*), there might be two different kinds of reasons for this kind of graphemic ambiguity. A phonetic form might be graphemically ambiguous

a. because the sound structure allows for the application of competing GPK rules (*mahlen-malen, Vetter-fetter, das-daß, Meer-mehr, du reihst-du reist*)

b. because phonemic distinctions, which are expressed graphemically, are removed by phonological rules (*Hund-Hunt, Tod-tot, (K)ranich-(t)ranig, du reist-du reißt*).

From a systematic point of view, there is a substantial difference between both types of ambiguity. Type (b) can be considered to be fully explained

within the presupposed framework. The different graphemic representations are derived in a completely normal way by the application of phonological rules and GPK rules. Type (a), by contrast, is based on the notion of markedness. At a certain stage of derivation, it is possible to apply two or more GPK rules to the same unit. In Bierwisch's approach, the choice between these rules is controlled by a system of markings. In a case like *Fuchs* vs. *Jux*, one form is considered as marked, i.e. not as the standard case. A GPK rule for these exceptions is only applied to those units which are marked for this rule. In our example, the marked GPK rule would convert /ks/ not to the normal <chs> but to <x> for marked units such as /juks/.

To use the concept of markedness is sometimes nothing more than a way of producing a correct solution by completely mechanical procedures. From the general theory of markedness, we know that the *real reasons* for units to be marked are of a very different nature. So we have to ask whether we are able to find out why certain units are marked with respect to the rules of the writing system and why others are not, and we will thereby find out what could be the reasons or some of the reasons for graphemic ambiguities. In what follows we will propose some preliminary answers to this question by showing how the choice between different GPK-rules is determined by the morphological properties of the units in question.

Example 1.

In our first example we are concerned with the relevance of inflexion vs. word formation (including composition) for writing. Even though it is not always easy to draw a sharp line between inflexion and derivation, this distinction is taken to be well-established in traditional and modern morphology. In most cases it appears to be irrelevant for the writing system, as can be seen from examples such as *Kinder-kind-kindlich* or *Könige-König-königlich*. As an example of where this distinction is relevant for spelling we will consider the phonological rule of geminate reduction in German (Geminatenvereinfachung), which was formulated and discussed in some detail in Wurzel 1970.

Roughly speaking, geminate reduction deletes certain phones in case they meet identical phones at a morpheme boundary. As Bierwisch has shown, this rule is relevant for writing. There are GPK rules which can only apply after the application of geminate reduction. So we write *du reist, du reißt, du reizt, du feixt* and not *du reisst, du reißst, du reizst, du feixst*. Now in Kohrt 1978 (63f.) it is argued that in some cases the GPK rules are working independently of geminate reduction. Whereas /t/ is deleted in forms such as *er rät* (from *raten*), it is not deleted in *er lädt* (from *laden*). In this

case the relevance or irrelevance of the rule simply appears to be determined morphonologically, but for other cases such as *er wird* from *werden* (not *er wirdt!*) this apparently does not hold. In general, there does not seem to be a simple criterion for the relevance of the rule and Kohrt states (ibid.) “daß die Grafemrelevanz bzw. -irrelevanz der Geminatenvereinfachung nicht durch einen bloßen Bezug auf bestimmte Klassen von Flexionsmorphemen gekennzeichnet werden kann.”

In the original formulation of geminate reduction by Wurzel, the necessity for this rule is demonstrated by the following list of phenomena (1970: 221f.):

- I. die ‘endungslosen’ Pluralformen bestimmter *ge*-Kollektiva (*die Gebirge, Gelege, Gestade*), deren zugrundeliegende Repräsentation /ge+STAMM+e+e/ ist;
- II. die Dativ-Plural-Formen der schwachen Nomina wie *den Buben, den Augen, den Straßen*, die aus /büb+n+n/ usw. abgeleitet sind;
- III. die Konjunktivformen wie *ich komme* und *du kommest*, die auf Repräsentationen /kom+e+e/ und [kom+e+est] (aus /kom+e+st/) zurückgehen;
- IV. Flexionsformen der 3rd. Person Singular wie z.B. *er rät* und *er brät*, in denen jeweils ein stammhaftes /t/ steckt (vgl. /rät+t/);
- V. Flexionsformen der 2nd. Person Singular wie z.B. *du heißt, du reißt, du mißt* (/rejs+st/ usw.).

It is immediately clear that in all these cases geminate reduction is relevant for writing. Furthermore these are all definitely cases of inflexion and not of derivation.

Wurzel then states (1970: 222), that geminate reductions ‘auch über die Begrenzungen von ‘orthografischen Wörtern’ (. . . vgl. *das ist tief* [istif]) und sogar von ‘phonologischen Wörtern’ im Sinne Chomskys und Halles (. . . vgl. *Schiff fahren* [šifar(ə)n] hinaus operieren. Ihre Domäne ist die phonologische Phrase (und damit manchmal der ganze Satz). “Although this is certainly correct for geminate reduction itself, it does not hold for its graphemic relevance. Instead I would like to claim that the maximal domain for GPK rules lies in general word forms.³ We will give another example later to support this claim. If it is correct, most cases of geminate reduction have no automatic consequences on the graphemic level. Furthermore it seems to be the case that geminate reduction is never relevant in cases of word formation, no matter whether these are cases of derivation (2a) or composition (2b).

- (2) a. *Schrifttum, unnachgiebig, ummodeln, annehmen, einnehmen, enttarnen*

b. *Stahllunge, Skatturnier, Tiefflieger, Waldessaum, heillos.*

It is not necessary to discuss the effect of the different forms of geminate reduction on spelling in order to be able to state that *all* cases of geminate reduction in writing are cases of inflexion. I would even claim that with inflexion, geminate reduction is *always* relevant for graphemics, at least if it has the effect of deleting one of two identical letters. If not, we get special rules which rely on special conditions, as in the case of *lädt*. Kohrt's *wird* instead of *wirdt* is not a real counter-example since the conjugation of *werden* is not completely regular anyway.

If our analysis is correct, then geminate reduction shows how different kinds of morphological facts can be directly relevant for spelling.

Example 2.

As is the case with word formation vs. inflexion, it is not always easy to decide whether a morphological unit or a morphological rule which makes use of that unit is productive, active, or neither, at any given time. Nevertheless, it is useful and even necessary that these terms in morphology exist simply because they refer to significant differentiations within the morphological system. Moreover, it seems to me that this distinction may also be of significance for the rules of spelling. We will illustrate this by referring to an interesting story told by Mark Aronoff and by complementing it by some further observations.

Aronoff, 1978, is concerned with the Latin suffix *-or*, its English and American descendents and their graphemic representations. In Latin, *-or* is used to nominalize intransitive verb stems, as in *error*, *tremor*, and for agent nominalizations on the basis of the supine, as in *ensor*, *victor*, *factor*. The counterparts of both suffixes *-or* were homographs in Middle English as well as in Anglo-French, where it was written as *-our*. In the sixteenth and seventeenth century, the etymologically correct form *or* was introduced into English. It then happened that the agent nominalizations were easily adapted to the new spelling, whereas the other type of nominalization was fairly resistant against it. That is why we now have *behaviour*, *rigour*, *labour*, *favour*, on the one hand, and *mediator*, *oppressor*, *supervisor* on the other. In American English this differentiation in spelling, which developed in English during the last three hundred years, was lost. The only written form was *-or*. Now there is also in English a suffix *-er* in addition to *-or* which has one function in common with *-or*, namely that of forming agent nominalizations from verbal stems such as in *writer*. Aronoff draws attention to the interesting fact that there is now a tendency in American English to write *all* agent nominalizations with *-er* and the other nouns with *-or*. Thus we find nomina

agentis, such as *adviser*, *sponser*, *adjuster*, but never **coler*, **behavier* etc. This would mean that within the system of American English, a pair of suffixes *-or*–*-er* is developing which is to some extent an analogue to the English *-ou*–*-or*. The morphologically and semantically well motivated differentiation in English is reintroduced into American English, yet with the effect of a simplified orthography.

In German we have *-or* (*Direktor*) as well as *-er* (*Programmierer*) and *-eur* (*Konstrukteur*) in agent nominalizations. For reasons of space, we will not discuss the distribution of these suffixes here.⁴ Instead we will elaborate a little on the use of a letter specific to German (and some other) system, namely the occurrence and nonoccurrence of ⟨ö⟩ ([ø]) in some classes of foreign words.

In foreign words, the French suffixes (*i*)*eux* and *eur* are written partly with ⟨ö⟩ and partly in the original way:

- (3) a. *muskulös*, *seriös*, *monströs*, *mirakulös*, *amourös*, *mysteriös*,
Likör
 b. *Ingenieur*, *Hypnotiseur*, *Dekorateur*, *Friseur*, *Masseur*.

It can be seen immediately that the difference in spelling must depend on the different status of the affixes *-ös* (*-iös*) on the one hand, and *-eur* on the other. The former suffix is not productive in German, in some cases it is not even recognized as a suffix at all. Even though the function of *-ös* as derivator of adjectival stems from noun stems is immediately recognizable in most cases, it is at least doubtful whether and in what way this regularity can be understood within the German morphological system. In many cases, the noun stems are not stems of German (not 'eingedeutscht'), as can be seen from examples like *muskulös*, *seriös*, *amourös*. These stems are not morphological units of the German system. That this construction is rather demotivated can also be seen from forms like *porös*. For many speakers, this word is not a derivative at all, but is a morphological simplex, in contrast to the form *porig* which uses the productive suffix *-ig*.

The situation is different with *-eur*, which first of all has preserved one of the meanings it has in French. With this meaning (deverbal nominalization referring to persons ('actors')) it is productive in German. The nouns with *-eur* are derived from verbs with *-ieren* (*massieren*–*Masseur*, *hypnotisieren*–*Hypnotiseur*, *frisieren*–*Friseur*) or from nouns with *-ion* (*Dekoration*–*Dekorateur*, *Konstruktion*–*Konstrukteur*) which in turn are derived from verbs with *-ieren*. This difference is probably motivated by phonetic facts (vowel or liquid in the terminal position of the stem in the latter cases). Furthermore, the nouns with *-eur* can themselves serve as the basis for derivations such as in *Masseuse*, *Friseuse*, *Ingenieurin*, *Dekorateurin*. Thus the

suffix *-eur* seems to be well integrated into the German system. I am taking this fact as the reason for the stableness of its written form. If a morphological unit is actively used within a language, and if there are no special reasons to change it (as in Aronoff's example), then it is reasonable to expect that this unit will be more likely to preserve its form than those units which are about to lose their morphological status. This could, for instance, explain why, despite many attempts to introduce the written form *Frisör*, we still write *Friseur*. On the other hand, we should write *Likör* and not *Likeur*, as we do, because this *-ör* has nothing to do with the *-eur*, neither morphologically nor semantically. Therefore [œ:] is written here as in most French loan words ⟨ö⟩ (*Möbel* etc.), even though there is a significant phonetic difference between the French form [œ:] and the German form [ø:]. This difference does not cause any problem because ⟨ö⟩ in German stands for both [œ] (*Spötter, Mörder*) and [ø] (*schön, möglich*).

Once again, this reflex of morphological facts in the writing system does not seem to have any counterpart at the sound level.

Example 3.

We will now consider in some detail the graphemic representation of the voiceless alveolar spirant [s]. As a separate letter or sequence of letters it is represented as ⟨s⟩, ⟨ss⟩, and ⟨ß⟩ in German. I am especially interested in finding out the rules which control the distribution of ⟨ß⟩.

The letter ⟨ß⟩ ([estset]) was introduced into the alphabet of German in the fourteenth century as a replacement of ⟨sz⟩. Its original function can roughly be understood as serving the purpose of avoiding the homography of two different [s]. To be more precise, ⟨ß⟩ was used as the graphemic counterpart of the [s] which originated from [t] by the Second Consonant Shift and then merged with the older [s] phonetically (cf. Müller 1978: 32f.). The role of ⟨ß⟩ in the writing system of present day German is not quite clear. There are many different proposals to formulate the distribution of ⟨ß⟩, but I know of none which comes to the conclusion that this distribution is at least approximately regular or even linguistically motivated. A typical statement to this effect is the one from a prominent orthography book by Mentrup which says that "In *many* cases there are reasons for choosing between ⟨s⟩ and ⟨ß⟩" (Mentrup 1968: 82; my translation, my italics). As a starting point for the analysis, we will present the rules for the distribution of ⟨ß⟩ according to Eichler 1978, 16ff. We have changed the presentation of the rules and translated them into English to make them easier to read, but we have not made any substantial change.

(4) [s] is represented graphemically

- a. as medial sound between long vowels and vowels always as ⟨ß⟩ (*rußen, Maße*)
- b. as medial sound between short vowels and vowels always as ⟨ss⟩ (*hassen, Nüsse*)
- c. as medial sound between vowels and consonants
 - α. as ⟨ß⟩ only if [s] also appears in the forms of the paradigm where the position in question is followed by a vowel (*spaßt–spaßen, paßt–passen*)
 - β. as ⟨s⟩ in all other cases (*Rast, Rost, Ast*)
- d. as medial sound between consonants as ⟨s⟩ if there are forms within the paradigm without [s] (*magst–mögen, hackst–hacken*)
- e. as final sound after vowels
 - α. as ⟨ß⟩ only if [s] also appears in those forms of the paradigm where the position in question is followed by a vowel (*Ruß–Rußes, Haß–Hasses*)
 - β. as ⟨s⟩
 - 1. if [z] appears in the same position (*Maus–Mäuse, Gas–Gases*)
 - 2. in foreign spellings (*Bus*, though we have *Busse*)
 - 3. in the suffix *nis* (*Finsternis*, although we have *Finsternisse*)
 - 4. exceptions (*aus, bis*)
- f. as final sound after consonants as ⟨s⟩
 - 1. if [z] appears in the same position (*Gans–Gänse*)
 - 2. as suffix *s* (genitive, plural) (*Abschieds, Muttis*)
 - 3. exceptions (*stets, bereits, allerdings*).

It should first be mentioned that this is one of the most complete and explicit statements of the rules under discussion. Its structure is imposed by the distinction between medial sound and final sound and by paradigmatic relations between word forms.

To reconstruct these rules in order to grasp more of the real structural conditions underlying the distribution of ⟨ß⟩, we first have to introduce the notion of morpheme boundary (#) and to understand the concept of medial sound as “not at the morpheme boundary”.

The only subrule from (4) which we are not going to change is (4b). We only reformulate the condition for the context. Instead of a short vowel, we speak of a vowel with the feature [–tense]. We do this because in other cases it is correct not to speak of long vowels but of tense vowels. So we will avoid speaking of long or short vowels altogether. The rule for ⟨ss⟩ from (4b) is now found as (5a). It seems to be important that [s] in German is *always* written as ⟨ss⟩ independently of any morphological conditions, in the context specified in (5a). That is to say that

(5) a.	{	⟨ss⟩ /	[−tense]	[+voc]		
b.		⟨s⟩ /	[+tense]	[+voc]/−#		
c. [s]→		⟨ʃ⟩ /	[+tense]	#St+	{	
d.		⟨s⟩ /	[−tense]	#St+		[−voc]Af
e.		⟨s⟩	else			∅

⟨ss⟩ is completely determined by phonetic conditions. It occurs both at morpheme boundaries (*hassen, Nüsse, gegossen*) and in the medial position (*Wasser, Kessel, Russe*).

Now ⟨ʃ⟩ behaves completely differently in this respect. With rule (5b) we first handle ⟨ʃ⟩ in the medial position. Our notation /−# says that a position *within a morpheme* is required. We understand the concept of morpheme in the Bloomfieldian sense as “minimal form”. We have to fulfill this condition if we want to refer to a medial sound in the sense stated above. We specify the preceding vowel as [+tense] because ⟨ʃ⟩ occurs after long vowels (*Straße, Muße*) as well as after diphthongs (*Meißen, perhaps außen*. I will come back to this example later).

With (5c) and (5d) we cover the occurrences of ⟨ʃ⟩ at morpheme boundaries. In (5d) it is specified that [s] is not written as ⟨ʃ⟩ at all morpheme boundaries but only in the final position of stems (St). The category St is a morphological constituent category. It is not possible to go into the pre-supposed morphological approach here (cf. Lieb 1977). I only want to mention that in this morphology, only three constituent categories are used, namely stem (St), affix (Af), and form (Fm). All other specifications such as stem “of a verb”, suffix “to derive nominals from verbs” etc. are not given in the constituent structure, but in what we call the marking structure. For the time being, it is only important that the constituent categories St and Af are sufficient to formulate the subrules (5c) and (5d): it is not important which kind of St or Af we have. By the notation St+ we mean “a stem which has always an [s] in final position”, as in *fließen, Maße*. Thus, St − is “a stem which can have [z] in final position” as in *reist, Maus, Gas*. Rule (5c) then states that [s] is written after tense vowels as ⟨s⟩ in the terminal position of stem morphemes, if [s] is preserved as terminal sound in all derivational and in flexional variants – in generative terms: if the underlying stem has an /s/ in final position. It can easily be shown that with (5c) we cover all cases from (4a) which are not covered by (5b), plus all cases from (4c, α) and (4e, α) with tense vowel, i.e. *spaßt, spaßen, rußen, Ruß* etc. The remaining cases are covered by (5d). This rule says that [s] is written as ⟨ʃ⟩ in the final position of stem morphemes if it is preceded by a lax vowel and if the following affix begins with a consonant, or if a suffix does not occur at all. The rule handles expressions such as *paßt, ist, Haß*.

There are some problematic cases with respect to (5), a few of which I should like to discuss very briefly. Forms such as *außer*, *außen*, *draußen*, *außerhalb* can be understood as being handled in two different ways by (5), depending on whether or not one wants to postulate a stem ⟨aus⟩. This is a difficult question, especially for *außer* and its derivatives. But we always get the correct spelling by (5), independently of the morphological analysis.

A different problem arises in connection with expressions such as *reist* from *reisen* as related to *reißt* from *reißen*. Here we have two different forms of geminate reduction. The contiguity of the final [s] and the suffix *-st* has different graphemic consequences. For *reist* it is again without consequence whether we consider the first or the second [s] as being deleted. We always get the correct spelling. But for *reißt* we have to postulate a stem ⟨rais⟩ because otherwise (5c) could not be applied. From the standpoint of morphology this is not an unreasonable commitment, I think.

As far as I know there exists only one expression which is not covered by (5). It is the conjunction [das], written ⟨daß⟩. Here we have an example of an effect of the so-called semantic principle of orthography by which in German homophone forms are sometimes discriminated graphemically.

What conclusions can be drawn about the writing system from a rule like (5)? First we have to state that – contrary to what our orthography books tell us – there seem to be practically no irregularities in the graphemic representation of [s]. This appears to be an important fact if one remembers that orthographic rules are the basis for teaching children the writing system. A linguistic analysis can prove that things are regular which normally are believed to be not at all regular or only partly regular.

Second, (5) seems to be much simpler than (4), although it is even more complete in the sense just indicated. This greater simplicity is quite substantial and is not primarily caused by the more restrictive language we have used to formulate the rules. There are other reasons for this simplicity.

First, we do not try to specify all cases positively as was done in (4). Instead we consider ⟨s⟩ to be the standard representation of [s] and we then specify the conditions for ⟨ss⟩ and ⟨ß⟩. By this we follow Bierwisch's proposal as to the difference between marked and unmarked cases. Yet we do not just mark some lexical entries for certain rules, but we specify the structural conditions for certain spellings.

As for the substance of (4) and (5) it seems to me that the main difference consists in our readiness to use morphological terms for formulating the rules.⁵ Since it would take too much space to comment on this in detail, I want to discuss only one point at some length.

Take the spelling of [us] in *Bus* (4eß2). This spelling is unusual according to (4) since the plural form *Busse* also has an [s], so we would have to write

Buß. The explanation would be that *Bus* is a foreign spelling. This explanation seems to be reasonable since indeed we have many words with *us* which are marked morphologically as 'foreign'. This can be concluded from the fact that these words have an unusual plural, i.e. a plural without Umlaut (see 6a), as we normally have it (see 6b).

- (6) a. *Omnibus–Omnibusse, Autobus, Fidibus, Habitus, Praktikus, Physikus, Krösus*
 b. *Kuß–Küsse, Schluß, Guß, Fluß, Nuß*

in (6a), *us* or *bus* are morphologically analyzed (and partly reanalyzed) as nominative singular affixes with the plural counterpart *usse*. So the forms in (6a) are certainly marked as foreign, but they are integrated into the morphological system in a way that differs from other words with *us* eg. those from scientific terminologies like *Kasus, Genus* etc.

This could be an argument for understanding the *us*-writing as 'foreign'. But, on the other hand, we also have many foreign words with an ⟨ß⟩, where ⟨ß⟩ is used in accordance with our orthographic rules. The forms in (7) show that ⟨ß⟩ is indeed used in foreign words if the structural conditions are met.

- (7) *Prozeß, Streß, Abszeß, Stewardesse, Narziß, Expreß*

In my opinion, the spelling of the nouns in (6a) is not irregular, but fits the system very well. We write *nisse–nis* exactly like *usse–us*, and this too is not an exception, according to the orthography books. Quite generally, we seem to have the simple regularity that [s] is never written as ⟨ß⟩ in an affix.⁶ One can easily check this by considering all so-called exceptions in (4). The distribution of ⟨ß⟩ is then determined by the following facts (stated partly in morphological terms).

- (8) a. in medial position, ⟨ß⟩ is determined phonetically
 b. in final position of stem morphemes, ⟨ß⟩ is determined morphologically
 c. ⟨ß⟩ normally occurs *only* in stem morphemes
 d. if ⟨ß⟩ occurs outside stem morphemes, there are special reasons for this ('semantic' principle for *daß*).

These results clearly indicate that certain regularities of our writing system are stated in the most simple and adequate way if one refers to morphological categories or, more generally, to genuinely morphological distinctions. As the greatest unit relevant for the writing system, we have word forms. All structural phenomena which go beyond word forms are irrelevant for the writing system, as far as spelling is concerned. Within word forms, we have to rely on morphological categories such as stem and affix. I am pretty sure that there

are many other rules of orthography which make reference to this distinction, e.g., rules for the doubling of consonants. Although this hypothesis cannot be confirmed at present, it should not be too surprising if it turned out to be true. With respect to their internal structure, affixes form a separate class of morphological units anyway, by what are called morpheme structure conditions, cf. Booij (1977: 22ff.)

Notes

1. With respect to German, consider for instance the collections of quotations from the literature given in Müller 1978 and Kohrt 1979.
2. In what follows, '[a]' denotes the phonetic segment *a*, '⟨a⟩' denotes the graphemic segment *a*, and '/a/' denotes the phonological segment *a* or simply 'the sound segment' *a* without specification of level of representation. Where no misinterpretations can arise, '⟨ ⟩' will often be omitted.
3. The word form as the domain for GPK rules can be exceeded if this is marked by a special character like the apostrophe. It is then possible that a rule for vowel reduction, which is a close relative of geminate reduction, is relevant for graphemics (see for instance French *Avenue d'Espagne, l'eau*).
4. With respect to spelling, things are completely different from English anyway since there is a clear phonetic difference in German between *-or* [o:ə] and *-er* [ɐ].
5. This seems also to be the main difference with respect to the most comprehensive study about the graphemic representation of [s] in German, which is found in Müller 1978. Like Fichler, Müller (1978:11) speaks of 'medial position', 'final position', etc., without referring to any other morphological units than word forms. Therefore, he is not able to systematically separate the distribution of ⟨s⟩ and ⟨ß⟩. Furthermore, Müller always states necessary conditions for the distribution of ⟨s⟩, ⟨ss⟩, and ⟨ß⟩, not sufficient ones. On the other hand, he gives some informal hints as to the relevance of morphosyntactic categories for the distribution of ⟨s⟩ (1978: 26).
6. Forms like *Steward*, *Baroneß* are no counter-examples to this regularity since *-eß* should not be considered as a German affix, despite the fact that we also have *Steward* and *Baron* (cf. Plank 1980:113ff.).

The form *Bus* has to be treated as analogous to the forms in (6a), though it certainly does not contain two morphemes.

References

- Aronoff, Mark
1976 *Word Formation in Generative Grammar*. Cambridge, Mass.: MIT Press.
- Aronoff, Mark
1978 "An English spelling convention." *Linguistic Inquiry* 9: 229-303.
- Bierwisch, Manfred
1972 "Schriftstruktur und Phonologie." *Probleme und Ergebnisse der Psychologie* 43: 21-44. Also in A. Hofer, ed., *Lesenlernen*. Düsseldorf: Schwann (1976), pp. 50-81.
- Booji, G.E.
1977 *Dutch Morphology. A Study of Word Formation in Generative Grammar*. Lisse: De Ridder.

- Chomsky, Noam and Halle, Morris
1968 *The Sound Pattern of English*. New York: Harper & Row.
- Eichler, Wolfgang
1978 *Rechtschreibung und Rechtschreibunterricht*. Königstein: Scriptor.
- Harris, William
1970 *Phono-Graphic Translation*. Manchester: Manchester University Press.
1975 'Writing: the basic options.' In W. Haas, ed., *Writing without Letters*. Manchester: Manchester University Press, pp. 131–208.
- Kohrt, Manfred
1978 Generative Phonologie und deutsche Orthografie. *Münstersches Logbuch zur Linguistik* 1: 49–76. (University of Münster).
- Kohrt, Manfred
1979 "Rechtschreibung und 'phonologisches Prinzip'." *Osnabrücker Beiträge zur Sprachtheorie* 13: 1–27. (University of Osnabrück).
- Lieb, Hans-Heinrich
1977 *Outline of Integrational Linguistics*. Fachbereich 16 der FU Berlin (= LAB 9).
- Mentrup, Wolfgang
1968 *Die Regeln der deutschen Rechtschreibung*. Mannheim: Bibliographisches Institut.
- Motsch, Wolfgang
1977 "Ein Plädoyer für die Beschreibung von Wortbildungen auf der Grundlage des Lexikons." In H. Brekle & D. Kastovsky, eds., *Perspektiven der Wortbildungsforschung*. Wuppertal: Bouvier, pp. 180–202.
- Müller, Rolf
1978 "Die s-Zeichen in der Linguistik und Didaktik der deutschen Schriftsprache." In *Laut und Schrift in Dialekt und Standardsprache*. Wiesbaden: Steiner, pp. 7–54 (= *Zeitschrift für Dialektologie und Linguistik*, Beiheft N.F. 27).
- Plank, Frans
1981 Morphologische (Ir-)Regularitäten. Aspekte der Wortstrukturtheorie. Tübingen: Narr. (Diss. Hannover 1980)
- Wurzel, Wolfgang Ulrich
1970 *Studien zur deutschen Lautstruktur*. Berlin (GDR) (= *Studia Grammatica VIII*).