

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/104908>

Please be advised that this information was generated on 2021-01-25 and may be subject to change.

The Interfaces of Writing and Grammar

1. Introduction

Spoken language, sign language, and written language – three modes of expression, but one underlying system? The answer will be negative for sign languages. Studies reveal that sign languages need not be derived from spoken languages and that if they happen to be derived from spoken languages, they tend to develop characteristics not present in their spoken origins (Wilbur 1987, Boyes Braem 1995). For younger generations, sign language can be acquired in a way that is familiar to how spoken languages are learned. Hence, there is evidence that sign language forms a system on a par with spoken language and is not dependent on it.

This is not the case for the written mode. Writing seems to be secondary to oral language, being derived from it, and fundamentally different from sign language. Each new generation learns the written variety at school after most of the spoken language has been acquired. Whereas for children the acquisition of a spoken or sign language is an unconscious process, acquisition of writing requires explicit learning strategies, of which teachers and pupils are well aware. Writing should be considered another code for the language acquired, which is why spelling is called secondary. The existence of spelling pronunciations, however, shows that this secondary mode of expression influences speaking, the primary mode (Van Haerigen 1962, Wells 1982: 106-9, Carney 1994, Maas 2000: 33). Other evidence for this influence on the primary mode comes from psycholinguistic experiments (cf., for instance, Seidenberg & Tanenhaus 1979, Schreuder et al. 1998) and from language change (Jespersen 1909). In this paper, the question how both modes of expression are related is investigated from a theoretical point of view.

The close relationship between a spoken language and its written variant has led to the hypothesis that the major part of the system is shared by both modes of expression. For instance, the semantic component provides the interpretation of scope-bearing elements, whether written or spoken; the syntactic component provides word order for both. Morphology creates words and inflection for both, and even some part of phonology is common, e.g. phonological segments correlate closely with letters. Some writing systems are called 'deeper' and others more 'shallow', reflecting the derivational level relevant for writing. Systems based on morphosyntactic structure are called deeper than systems based on phonological or phonetic representations (Haas 1976, Sampson 1985, Sgall 1987, Asher & Simpson 1994, Daniels & Bright 1996, Meisenburg 1996). The claim is that the written mode of expression follows a route different from the oral mode only in the final stage of processing. In reading, it is only the first stage of processing that follows a different route, according to this hypothesis. Speaking and writing thus share a large number of derivational stages, as do hearing and reading. Schematically:

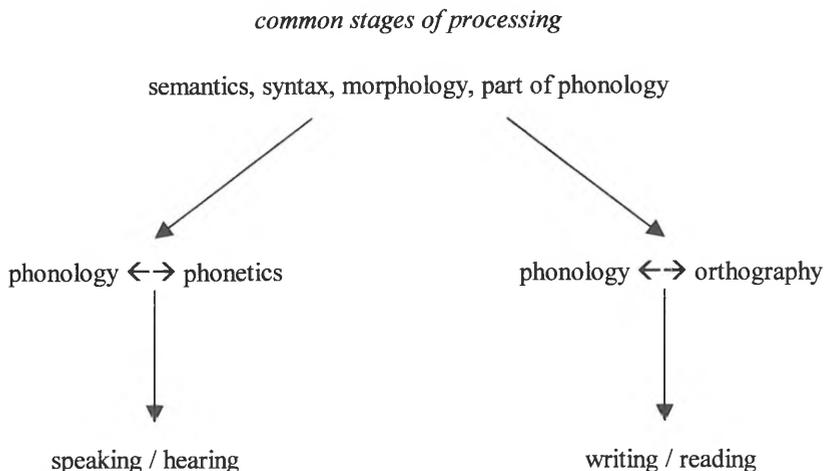


Figure 1: *General model of the relation between spoken and written language*

This view on how spoken and written language relate to each other has been worked out for Dutch by Nunn (1998). Dutch orthography is known to be based on a deep phonological stage of processing, cf. Van Heuven (1978) and Booij (1987). Nunn (1998) adds to this the conclusion that the derivation from phonology to orthography consists of two steps. After the first step of phoneme-to-grapheme conversion for morphemes, a second step takes care of grapheme co-occurrence restrictions by way of graphotactic rules, i.e. grapheme-to-grapheme conversion rules. Nunn calls such rules ‘autonomous spelling rules’, claiming that the rules refer to orthographic information only, although some of the phonological characteristics (the distinction between consonants and vowels, for instance) are carried over to the orthographical representation.

Of course, in defending the claim of a derivation in two steps, Nunn emphasizes the differences between the two steps, i.e. the difference between phonologically and orthographically based rules. It is from this perspective that Nunn tries to find evidence for the orthographic nature of autonomous spelling rules and to restrict the amount of phonemic information necessary for the second step in the derivation from phonology to writing. From this perspective, it is not surprising that Nunn’s analysis of Dutch has been used in Sproat (2000: 16) to illustrate the Consistency Hypothesis.

(1) *Consistency*

The Orthographically Relevant Level for a given writing system (as used for a particular language) represents a consistent level of linguistic representation.

This hypothesis, a direct reflection and strict interpretation of the model sketched in figure 1, states that there is one consistent Orthographically Relevant Level for a given writing system, not more than one, cf. figure 2. Notice that ‘Consistent’ here must not be understood as ‘without exceptions’. Where alphabetic writing systems concern the spelling of finite sets of elements, the opportunity is present to store exceptional orthographic forms in memory. It seems that exceptions occur in many alphabetic writing systems.

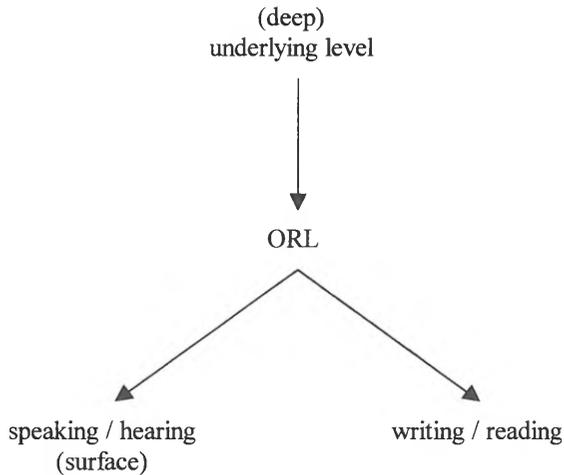


Figure 2: *The claims of the Consistency Hypothesis: one consistent level by the oral and written modes*

In this paper, evidence will be presented to show that the processes of speaking and writing share more information than can be provided by a single derivational level. The claim made in this paper is that the phoneme-to-grapheme conversion rules are based on information from different levels, as are the grapheme-to-grapheme conversion rules. Of course, the distinction between the two sets of rules will be valid even when more than just one linguistic level provides input to the orthographic representation. Therefore, the two-step analysis of Nunn can be maintained, though defined in a less rigorous fashion. The Consistency Hypothesis, however, cannot be maintained as a universal principle.

The layout of this paper is as follows. First, the arguments by Nunn (1998) in favor of a two-step derivation of orthography will be reviewed. Then, in section 3, the Orthographically Relevant Level according to Nunn will be discussed. It will be shown that the hypothesis that there is only one such level can be maintained only at the cost of storage. Sections 4 and 5 show that a native Orthographically Relevant Level must be distinguished from a non-native Orthographically Relevant Level and that punctuation is based on other levels than the phonemic representation of morphemes. Section 6 presents the linguistic information necessary for the autonomous spelling rules. Section 7 finally summarizes the evidence gathered in the preceding sections about the linguistic levels needed for writing and presents the overall conclusion. Information from different levels of language processing is collected in writing. In the presentation that follows, most arguments are based on writing and virtually no arguments are presented about reading.

2. An outline of Dutch orthography

Detailed information on the orthography of Dutch can be found in Nunn (1998). She distinguishes several orthographic components that are relevant for Dutch. Conversion of native morphemes needs to be distinguished from conversion of non-native morphemes, and a set of autonomous rules forms part of the orthographic derivation. Figure 3 is Nunn's analysis in a nutshell. Observe that she assumes one level with information on the underlying, phonemic, representations of the segments of morphemes at which all information of the spoken mode is translated into information on the written mode. Nunn's proposal for Dutch therefore confirms Sproat's Consistency Hypothesis. According to Nunn, there is one Orthographically Relevant Level, the level of morphemes in their phonemic form:

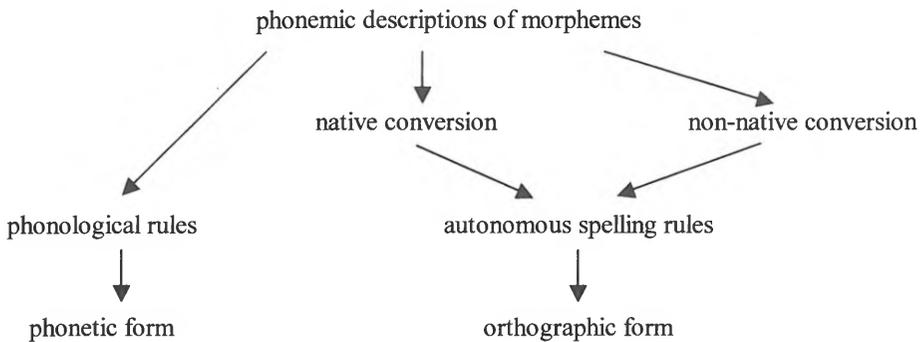


Figure 3: Nunn's model of the relation between phonetic and orthographic form

The remainder of this section will present explanatory notes on this model.

Dutch has a so-called deep orthography. Underlying rather than superficial sound segments are spelled; i.e., morphemes tend to receive a uniform spelling, irrespective of the application of certain phonological rules that generate sets of allomorphs. Frequently used examples to illustrate this are *hond* and *heb*, with final obstruents spelled in accordance with their underlying forms /hɔnd/ and /heb/ instead of their phonetic forms [hɔnt] and [hep]. These underlying forms are detectable for the writer on the basis of plural inflection: [hɔndə] and [hebə] with voiced obstruents. Other examples are *zuinigheid*, *aanmelden*, *hoofdoek* 'carefulness, to announce, head-shawl', for which a more superficial spelling would be **zuinigeit*, **aamelde*, **hoofd**oek*, derived by h-deletion, final devoicing, nasal assimilation, final n-deletion, and degemination.

Furthermore, Dutch is a language with two sets of words: native ones, such as *kunstzinnigheid*, and non-native ones, such as *artisticeit*, both meaning 'artisticity'. The difference has its origins in the earlier stages at which Dutch imported words from Latin or French, but new borrowings follow this distinction as well. Non-native words can be distinguished from native ones on the basis of systematic differences in present-day phonology and morphology (Van Heuven et al. 1994, Nunn 1998: 155 ff.). One of the most important characteristics is the

number of full vowels present in morphemes: when more than one full vowel is present, the morpheme will be non-native. Exceptions to this rule are only a handful of frozen compounds such as *aardbei* 'strawberry' which behave as native words, notwithstanding the presence of more than one full vowel.

The distinction between native and non-native morphemes takes the native morphemes as point of departure, such that all morphemes not in accordance with the constraints that hold for native morphemes are non-native. Therefore, the fact that only one full vowel is present in a morpheme is a necessary but not a sufficient criterion for this morpheme being a native morpheme. Further constraints are the combination of consonant clusters (for instance, only a limited set of clusters occurs in native morphemes, not the clusters /sk/, /sf/, and /tm/, which predicts that *skelet* 'skeleton', *sfeer* 'sphere', and *ritme* 'rhythm' are non-native words, even though only one full vowel occurs) and constraints on morphology (for instance: plural -s is restricted to native words ending in /a, o, u/ and native words ending in a syllable with schwa; hence, the plural forms *trams* and *e-mails* indicate that these words are non-native). On the basis of such criteria, the etymological distinctions are recoverable from the synchronic spoken mode even for language users without any knowledge of foreign languages.

The orthography reflects the difference between native and non-native words, since partly different sets of phoneme-to-grapheme conversion rules are used (indicated in figure 3 by the two routes for native and non-native morphemes) with, for instance, the graphemes *c, q, th, y,* and *x* for non-native words only, cf.:

(2)	<i>non-native</i>	<i>native</i>	<i>sounds</i>
	<u>camera</u> 'camera'	<u>kamer</u> 'room'	/k/
	<u>quasi</u> 'quasi'	<u>kwaad</u> 'angry'	/k/
	<u>ether</u> 'ether'	<u>eter</u> 'eater'	/t/
	<u>hypothese</u> 'hypothesis'	<u>hier</u> 'here'	/i/
	<u>examen</u> 'test'	<u>heks</u> 'witch'	/ks/

Literacy therefore leads to awareness of the distinction between native and non-native morphemes.

The general model in figure 1 of how speaking and writing can be related is not only complicated by the difference between the spelling of native and non-native words, but also by the existence of autonomous spelling rules. One of the reasons to incorporate such rules in the model of Dutch orthography is the presence of allography in examples such as:

(3)	<i>stem</i>	<i>derived form</i>	<i>spelling</i>
	bak	- bak+er →	bakker 'baker'
	judo	- hij judo+t →	hij judoot (third person ending of the verbal stem <i>to judo</i>)
	laan	- laan+en →	lanen 'lanes'
	vers	- iets vers+s →	iets vers 'something fresh'

No phonological alternation is involved here. In order to account for such forms of allography, Nunn (1998: 183 ff.) proposes a set of autonomous graphotactic rules, i.e. rules that operate on grapheme sequences, such as the following ones for gemination and degemination. The formulation of Nunn's rules has been simplified for expository reasons. C abbreviates for consonant letters, V for vowel letters, and dots indicate syllable boundaries. The distinction between short and long vowels is not one of phonetic duration, but rather expresses the fact

that short vowels may combine with a coda that consists of more consonants than the coda following long vowels.

- (4) a. *Orthographic gemination*
 C → CC after a short vowel at the end of the syllable
 V → VV for long vowels when a C follows within the syllable
- b. *Orthographic degemination*
 VV → V when syllable final
 CC → C when syllable final

The derivation of the words presented in (4) runs as follows (backslashes indicating the underlying orthographic forms):

- (5) a. *Conversion of morphemes*
 /bak/ → \bak\ /ər/ → \er\
 /jydo/ → \judo\
 /lan/ → \laan\
 /vers/ → \vers\ /s/ → \s\
 /t/ → \t\
 /ən/ → \en\
 /s/ → \s\
 b. *Concatenation of morphemes and syllabification*
 \ba.ker\
 \ju.dot\
 \laa.nen\
 \verss\
 c. *Application of orthographic (de)gemination rules, cf. (4)*
 <bakker>
 <judoot>
 <lanen>
 <vers>

As a result of these orthographic rules, vowel letters for short vowels are always followed by a consonant within the syllable, whereas syllable-final vowel letters represent long vowels. It is because of this pattern that short and long vowels in the literature on Dutch orthography are called 'covered vowels' and 'free/uncovered vowels' (Dutch *gedekte* and *ongedekte/vrije vocalen*). Covered vowels are always followed by a consonant letter within the syllable, whereas uncovered vowels may occur at the end of syllables:

- (6) *covered/short vowels* *covered by C-gemination* *uncovered/long vowels*
 [kantə] kan.ten 'sides' [kɑ.nə] kan.nen 'cans' [manə] ma.nen 'moons'
 [kældər] kel.der 'cellar' [bɛ.lə] bel.len 'bells' [benə] bɛ.nen 'legs'
 [plɔftə] plof.te 'plumped' [plɔ.fə] plof.fen 'to plump' [pokər] po.ker 'poker'

This generalization holds in orthography but is present in phonology as well: intervocalic consonants after short vowels are ambisyllabic, as demonstrated in experiments in which speakers of Dutch are forced to explicitly syllabify such examples (cf. Rietveld 1983 and Sandra et al. 1996). The experiments show that speakers' judgments are influenced by orthography. However, interestingly, illiterate speakers of Dutch and pre-school children also present analyses with ambisyllabic consonants, though significantly less than the literate participants for whom the spelling rules seem to enhance ambisyllabic responses.

The only exceptions to the generalization that short/covered vowels are followed by a consonant within the syllable are loan words such as *sjwa* [sjwa] ‘schwa’ and exclamations such as *bah* [ba] ‘ugh!’ and *joh* [jo] (an exhortative word, presumably derived from *jongen* ‘boy’). In the latter cases, <h> has the function of covering the short vowel at the level of orthography.

With the introduction of autonomous orthographic rules, the process of writing becomes a two-step derivation. The first step is conversion from phonemes to graphemes and the second step is the set of autonomous spelling rules for the conversion from graphemes to graphemes. Arguments in favor of this position are based on the observation that the two sets of rules display different characteristics (Nunn 1998: 131):

(7)	<i>phoneme-to-grapheme</i>	<i>autonomous</i>
	<i>conversion rules</i>	<i>spelling rules</i>
<i>context</i>	phonological	orthographic
<i>domain</i>	morpheme	word
<i>native/non-native sensitive</i>	yes	no

The following short summary of spelling /i/ in Dutch will illustrate the characteristics of the conversion rules (backslashes again indicate underlying orthographical forms):

- (8) *Conversion rules for /i/*
 a. /i/ → \ie\ in native morphemes (*kietel* ‘kittle’)
 b. /i/ → \ie\ in the last syllable of non-native morphemes (*komiek* ‘comic’, *natie* ‘nation’)
 c. /i/ → \i\ in non-native morphemes, when not the last syllable (*titel* ‘title’)

These rules take phonological information as their input and are restricted to the morpheme-domain. Rules (a) and (c) show that the native/non-native distinction is relevant. Both *kietel* and *titel* are monomorphemes, and hence, only the conversion rules can be responsible for the spelling difference. Rule (b) shows that information on morpheme boundaries is essential. Final syllables in native and non-native morphemes are spelled <ie>.

Diaeresis placement may illustrate the characteristics of autonomous rules. This rule applies to ambiguous letter strings: because *aa*, *oo*, and *uu* encode one sound in (9a), a diaeresis should be used in (9b) where the two vowels indicate two sounds (resp., uncovered/long and covered/short ones). Because *ii*, *eo*, and *ue* are not in use as a digraph, no diaeresis should be used for these letter pairs, cf. (10).

- | | | |
|-----|---------------------------------|----------------------------------------|
| (9) | a. [a] <u>baal</u> ‘bale’ | b. [aɔ] <u>Baäl</u> ‘biblical name’ |
| | [o] <u>koor</u> ‘choir’ | [oɔ] <u>coördinatie</u> ‘coordination’ |
| | [y] <u>postuum</u> ‘posthumous’ | [yʏ] <u>vacuum</u> ‘vacuum’ |
-
- | | | |
|------|----------------------------------------------|---------------------|
| (10) | a. [iɪ] <u>kopiïst</u> ‘copyist’ | b. * <u>kopiïst</u> |
| | [əɔ] <u>geolied</u> ‘oiled’ | * <u>geölied</u> |
| | [yə] <u>ambiguë</u> ‘ambiguous + inflection’ | * <u>ambiguë</u> |

The following examples show that diaereses occur in native and non-native forms alike:

- | | | |
|------|------------------------|-----------------------------|
| (11) | <i>native words</i> | <i>non-native words</i> |
| | geënt 'grafted' | geëmotioneerd 'emotional' |
| | knieën 'knees' | manieën 'manias' |
| | beïnvloed 'influenced' | geïllustreerd 'illustrated' |

Moreover, these examples are morphologically complex, which shows that the diaeresis rule also applies across morphological boundaries, at the level of the word. Diaeresis placement will be discussed below in section 6.1, where the fact that the rule makes use of phonological information will lead to the conclusion that such graphotactic rules are not autonomous.

In sum: Nunn finds evidence for her two-step hypothesis in the clustering of characteristics of the rules involved. Her two-step analysis will be taken as a point of departure for the remainder of this paper, but arguments will be presented against the claim that the context of autonomous spelling rules consists purely of orthographic information. First, evidence will be presented that phoneme-to-grapheme conversion rules are based on information from different levels (section 3) and that the Orthographically Relevant Level is different for native and non-native words (section 4).

3. Phonological rules expressed in Dutch writing

In this section, rules will be discussed that show that some morphemes are spelled according to the phonemic level, but that a more superficial level must be assumed for other morphemes. Nunn's conclusion was that the more superficially spelled allomorphs are stored in the lexicon, even though phonological rules predict their distribution. In the absence of independent arguments for this position, one might claim equally well that the cases discussed are counterexamples to the Consistency Hypothesis and that there are several Orthographically Relevant Levels for Dutch.

3.1. Voice assimilation

As illustrated above, Dutch spelling is based on a deep phonological level at which, for instance, the rule of Final Devoicing has not been applied. *Hond* and *heb* are written, even though [hɔnt] and [hɛp] are pronounced. Dutch orthography, however, reflects Perseverative Devoicing in past tense suffixes, cf.:

- | | | |
|------|----------------------------|---------------------|
| (12) | <i>Dutch past tenses</i> | |
| | [d] stem – stem <u>d</u> e | 'vote – voted' |
| | tob – tob <u>d</u> e | 'worry – worried' |
| | kano – kano <u>d</u> e | 'canoe – canoed' |
| | [t] lek – lek <u>t</u> e | 'leak – leaked' |
| | hoop – hoop <u>t</u> e | 'hope – hoped' |
| | straf – straf <u>t</u> e | 'punish – punished' |

The fact that Perseverative Devoicing in past tenses is expressed in spelling comes as a surprise, given that Dutch spelling generally expresses the underlying form of d/t-allophony.

The inconsistent spelling of past tenses has been explained in the literature by referring to Readability, an output constraint that requires *lekte*, *hoopte*, and *plofte* instead of **lekde*, **hoopde*, and **plofde*. The Readability Requirement has been incorporated in the Principle of Uniformity by Te Winkel (1863: 12) as follows:

- (13) *Principle of Uniformity* (Regel der Gelijkvormigheid)
Give the same orthographic form to a word and to its constituent parts, *as far as pronunciation allows this*.¹

This explanation has been repeated in later publications (for instance, in Booij et al. 1979), but the Readability condition has never been explicitly formulated (but cf. Neef, this volume).

There are reasons to doubt that Readability can be so formulated that it accounts for the spelling of past tenses in Dutch. First, observe that the reading process is quite robust, as illustrated by examples such as *politie* 'police' and *politiek* 'politics' (with <tie> indicating [tsi] in the first word and [ti] in the second one) and diminutives such as *crèmepeje* (lit. 'small cream', i.e. cream in small pots or tubes), written with three syllables and pronounced with only two. Such examples show that the relation between spelling and pronunciation may be a loose one, as long as the morphemes are recognized and get a stable spelling.

The second argument comes from English. Observe that past tenses in English are also subject to Perseverative Devoicing, but that these verbs receive a morphological spelling.

- (14) *English past tenses*
[rd] lift – lifted
[d] puzzle – puzzled
[t] look – looked

If these forms are not problematic for English readers, why then would the deep spellings **lekde*, **hoopde*, and **strafde* be problematic for Dutch readers? Presumably, Readability is a universal requirement, related to the language processing capacities available to human beings. When languages differ, the differences should be explainable on the basis of other characteristics of the languages, and no such explanation seems to be available for these cases.

In order to maintain the hypothesis that phonemic representations of morphemes form the input for spelling, Nunn proposes a lexical approach to past tense allophony. She assumes that these suffixes are stored in their more superficial forms *-te* and *-de* (cf. Nunn 1998: 63 and 136) and that not only storage of underlying forms, but also the option of what she calls 'competing allomorphs' is available in Dutch orthography. Evidence from the spoken mode for the special status of past tense suffixes is then called for. As long as such evidence is lacking, these instances might as well illustrate that some morphemes (be it a finite list) get a more superficial spelling, whereas the spelling of most morphemes is in agreement with the underlying phonological representation. But this alternative would be in conflict with the Consistency Hypothesis.

¹ Geef, zooveel de uitspraak toelaat, aan een zelfde woord en aan ieder deel, waaruit het bestaat, steeds denzelfde vorm.

Another way to maintain the Consistency Hypothesis, more in line with Sproat (2000), would be to assume that Dutch orthography is based on some intermediate level, after the application of Perseverative Devoicing but before all other phonological rules apply. This proposal conflicts with the traditional, derivational approach of Dutch voicing assimilation present in the literature on Dutch phonology; cf. Zonneveld (1983), who claims that Final Devoicing is ordered before all other assimilation rules. Of course, other analyses of Dutch voicing assimilation can be provided. For instance, analyses with another domain of application for Final Devoicing (not the word, but the syllable), with another underlying form of the past tense suffix, or with a lexically governed rule of Perseverative Devoicing, different from the general rule of Progressive Assimilation in Dutch. But it seems hard to find independent evidence to choose between these alternative approaches, which is one of the reasons why the derivational approach is no longer the predominant model of phonological research.

As long as derivational models do not provide independent evidence for a specific rule ordering, the conclusion must be that the Consistency Hypothesis cannot be tested. Rather than forwarding claims about some intermediate level and more in line with newer insights in the interaction of components, one should conclude that the underlying phonemic representation is the input to Dutch orthography, except for a finite list of morphemes of which the allomorphs are distinguished in orthography.

3.2. Diminutive allomorphy and d-insertion

Nunn (1998: 62) proposes the competing allomorph analysis also for the spelling of diminutives and for the spelling of agentive and comparative *-er*, cf.:

(15) *allomorphy/allography of diminutive suffixes -etje, -tje, -pje, -kje, -je*

bloem – bloemetje	‘flower’
laan – laantje	‘lane’
oom – oompje	‘uncle’
koning – koninkje	‘king’
koek – koekje	‘cookie’

(16) *allomorphy/allography of -er, -der*

roep – roeper	‘call – caller’
hoor – hoorder	‘hear – hearer’
mooi – mooier	‘beautiful – more beautiful’
raar – raarder	‘weird – weirder’

The choice between diminutive suffixes is predictable on the basis of phonological contexts, but some diminutives get an idiosyncratic meaning which may form an argument for considering diminutive allomorphy as a lexicalized process. The rule of d-insertion before agentive and comparative *-er*, on the other hand, is productive and fully predictable. Therefore, the competing allomorph analysis is not more likely to be present for these morphemes than it is for any other morpheme. Hence, as long as no additional evidence is provided, *-er/-der* allography forms an argument against the Consistency Hypothesis.

full answer. Observe that according to this hypothesis, other instances of Dutch orthography will be inconsistent. For instance: devoicing of fricatives at the end of words is a general rule of Dutch phonology, but still, the superficial form is spelled in words such as *huis* and *leef* (which have underlying voiced fricatives, witnessed by the inflected forms *huizen* and *leven*). Other examples illustrating that there is no tendency to avoid representation of general phonological rules in Dutch are vowel reduction in words such as *apostel*, *cirkel* ('apostle, circle', with derived forms *apostolisch* and *circulair*), nasal assimilation in monomorphemic words such as *ramp* 'disaster', and degemination at the end of words. Perhaps all these counterexamples can be explained on the basis of graphotactic rules, but the question to be answered then is why such graphotactic rules violate an otherwise sensible constraint on the orthographic system for Dutch.

4. Native and non-native morphology

The spelling of non-native words in Dutch differs systematically from the spelling of native words. Above, in (2), examples are presented with *c*, *q*, *th*, *y*, and *x*, letters that are not in use for the sounds /k, t, i, ks/ in native words. More subtle differences exist in the spelling of vowels. In native words, long (or uncovered) vowels are spelled with digraphs and short (or covered) vowels are written with a single letter, cf. (20). In non-native words, however, all vowels are written with a single letter, cf. (21a), except when they occur in the final syllable of the word, cf. (21b):

(20) *native words*

<i>short vowel</i>	<i>long vowel</i>
[ɑ] <u>h</u> andel	[a] va <u>a</u> ndel
[ɛ] <u>v</u> erder	[e] me <u>e</u> erder
[ɪ] <u>m</u> ispel	[i] <u>k</u> ietel
[ɔ] <u>k</u> oster	[o] <u>k</u> looster
[ʏ] <u>d</u> urven	[y] <u>h</u> uurder

(21) a. *non-native words, nonfinal syllables*

<i>short vowel</i>	<i>long vowel</i>
[ɑ] <u>a</u> potheek	[a] <u>a</u> potheose
[ɛ] <u>e</u> cho	[e] me <u>c</u> hanisch
[ɪ] <u>d</u> ijstichon	[i] <u>d</u> iploma
[ɔ] <u>c</u> omité	[o] <u>h</u> omoniem
[ʏ] <u>m</u> uskiet	[y] <u>s</u> tucwerk

b. *non-native words, final syllables*

<i>short vowel</i>	<i>long vowel</i>
[ɑ] <u>s</u> esam	[a] amalga <u>a</u> m
[ɛ] <u>r</u> e <u>b</u> el	[e] ga <u>r</u> e <u>e</u> l
[ɪ] <u>p</u> assim	[i] muzi <u>e</u> k
[ɔ] <u>c</u> omplot	[o] <u>p</u> iloot
[ʏ] <u>m</u> useum	[y] <u>m</u> in <u>u</u> ut

The difference between long and short vowels does not seem to be a phonemic difference in non-native words, which is why variation of pronunciation may occur. For instance: *apothek* with a first long vowel and *apotheose* with a first short vowel occur, though perhaps less frequently as the other way around. Only a keen listener will notice when *musket* and *stucwerk* are pronounced [myskit] and [strykwerk] instead of the more usual [myskit] and [strykwerk].

Minimal pairs based on vowel length (such as *komma* ‘comma’ and *coma*) are hard to find in the set of non-native words, presumably because vowel length distinctions played a minor role in the donor language Latin, in which *liber* ‘book’ and *liber* ‘free’ is one of the few examples of a minimal pair based on this distinction.

The above examples illustrate that different conversion rules apply to the two classes of words. The following examples show that the domain at which conversion takes place differs also (Nunn 1998: 93):

(22)	a. <i>stem</i>	b. <i>native suffix</i>	c. <i>non-native suffix</i>	
	Fries	Friezin	frisisme	‘Frisian – Frisian woman – frisism’
	limiet	limieten	limiteer	‘limit – limits – to limit’
	trochee	trocheeën	trocheïsch	‘trochee – trochees – trochaic’
	station	stationnetje	stationair	‘station – small station – stationary’

Starting with a native or non-native stem, cf. (22a), a native morpheme added to it results in a spelling without adaptation, cf. (22b). This is what the Consistency Hypothesis predicts in combination with the assumption that morphemes form the domain of phoneme-to-grapheme conversion. When, however, a non-native suffix is added, the stem is spelled as if the word were monomorphemic: long vowels are written with a single letter, cf. (22c). Nunn accounts for this spelling behavior by assuming that non-native morphology is ignored. Complex derivations with non-native affixation are treated as if they were monomorphemic.

The solution proposed by Nunn meets some difficulties. First, the above examples of contraction and emphatic use presented in connection with Nasal Assimilation (*in-en export*, *in-plausible* etc.) show that morphological structure is present in non-native derivations. Second, a set of correspondence rules is needed to account for spelling idiosyncrasies that occur in non-native sets of words such as *context* – *contextueel*, *tekst* – *intertekst* – *intertekstueel*, *medievist* – *medievistiek*, *quaestor* – *quaestrix*. Morphemes of non-native complex words receive a constant, though sometimes idiosyncratic spelling, but the spelling of sets of morphologically related non-native words cannot be considered completely ad hoc. *Context* and *tekst* form the basis of the two sets of consistent spellings; *ae* is replaced by *e* in *ether* (< *aether*) and forms derived from *ether*, but not in *quaestor* and its derived forms. Third, consonant geminates in non-native words are the reflection of morphological structure, cf. *acclamatie* – *declamatie*, *adduceren* – *deduceren*, *collocatie* – *dislocatie*. When writers are aware of this kind of morphology, this shows that non-native morphology is present in the language system and reflected in orthography.

On the other hand, some distributional facts will receive an explanation by a level in between non-native and native morphology (cf. Van Beurden 1987): given such a level, non-native morphemes would be closer to roots and stems than native morphemes, which is generally true, although productive formations to the contrary exist (cf. Haas & Trommelen 1993: 459 ff.) (in (23), boldfaced *sub-*, *hyper-*, and *-eer* are non-native affixes, the other affixes are native):

- (23) a. *General pattern*
- | | | |
|-----------------|--------------------------------|-------------|
| <i>spelling</i> | <i>morphology</i> | |
| groepering | ((((groep) <u>eer</u>) ing) | 'grouping' |
| verdisconteer | (ver((dis(cont)) <u>eer</u>)) | 'negotiate' |
- b. *Exceptions*
- | | | |
|-----------------|-----------------------|-------------------|
| <i>spelling</i> | <i>morphology</i> | |
| subafdeling | (sub(af(deel))ing) | 'subsection' |
| hypergevoelig | (hyper((ge(voel))ig)) | 'hyper-sensitive' |

Perhaps *sub-*, *hyper-*, and the like are to be grouped together with the native ones. (For English, non-native prefixes are claimed to belong to Class II, cf. Giegerich 1999 and previous literature.) In that case, the different spelling behavior of native and non-native morphemes can be combined with the Consistency Hypothesis when a level in between non-native and native morphology is assumed to form the input for phoneme-to-grapheme conversion, and the elements converted are native morphemes and non-native complex forms. In that case, a new solution must be found for idiosyncratic spellings of related non-native formations and for consonant geminates in contexts where orthographic gemination does not apply.

The level ordering hypothesis and stratum-oriented models never succeeded in adequately describing the morphological patterns available in languages such as English, German, and Dutch. Instead, approaches with restrictions for individual morphemes seem to be more successful, cf. Fabb 1988, Neef 1996, Plag 1999, and Hay 1999. In line with these more recent approaches, the competing allomorph analysis forwarded by Nunn, and hence storage of spelling forms for individual morphemes, seems to be more promising than the search for one level as the input for writing.

5. Punctuation

The orthographic rules that mirror segmental phonology take morphological words as their maximal domain. For other aspects of orthography, i.e. punctuation, larger domains are relevant. For instance: words in a phrase are separated by spaces, words in compounds are written together. When a phrase is embedded in a word, the spaces are eliminated, which offers the opportunity to disambiguate in writing what may be ambiguous in speaking, cf.:

- | | |
|----------------------------------------------|------------------------------------------------|
| (24) <i>phrase</i> | <i>compound</i> |
| klein kind 'small child' | kleinkind 'grandchild' |
| vuile grondaffaire 'dirty affair about land' | vuilegrondaffaire 'affair about polluted land' |
| oude mannen 'old men' | oudemannenhuis 'old men's house' |

Moreover, larger syntactic or prosodic constituents can be distinguished in orthography. Capital letters and dots surround utterances, and commas or semicolons separate the parts of enumeration (as in English), with subtle distributional differences that may signal coordination embedded within coordination: *Jan, Karel en Harry; Kees, Marie en Piet; en Susan, John en Martin*. There are indications that the distribution of punctuation signs relates to the

thematic structure of the text (cf. Bredel, this volume). The following list of orthographic means for representing information from different linguistic levels illustrates the issue, but is not meant to be exhaustive:

(25)	<i>orthography</i>	<i>linguistic level</i>
	spaces	syntax: syntactic words
	capital letters	semantics: proper names
	capital letters	syntax: German nouns
	capital letters and dots	syntax or prosody: utterances
	commas	syntax: clauses, coordinated constituents
	semicolons	syntax: coordination with embedded coordination
	indents or white lines	syntax: domain of pronominal reference
	dash	semantics: change of focus

This list shows that the Orthographically Relevant Level needs to be an all-encompassing representation of the utterance, including syntactic and semantic information. The Consistency Hypothesis claims, however, that writing is based on a single level of information, e.g. that it translates sounds of a certain level into letters.

6. Linguistic information for graphotactic rules

According to the null hypothesis, the output of phoneme-to-grapheme conversion would be the string of letters. Nothing more. However, there is abundant evidence in the formulation of autonomous spelling rules that the representation is more articulate. Word and morpheme boundaries are retained, as is some information about the connection between successive words and morphemes, witnessed by the use of spaces, for instance. In this section, the kind of information necessary in the formulation of graphotactic rules will be discussed. It will be shown that graphotactic rules and phonological rules of the later stages of the derivation share characteristics which make it necessary to assume a close relationship between the two sets of rules.

The discussion will take as its point of departure the proposal by Nunn (1998: 32-34) to carry over to orthography the morphological structure and specific parts of the segmental phonological information, e.g. the distinction between vowel letters and consonant letters and information about length of vowels. Nunn proposes representations such as (26) to handle the fact that <ee> and <e> may be long vowels (VV) and that <e> may represent a short vowel (V) as well.

(26)	<i>letter tier</i>	ee	e	e
			^	
	<i>CV-tier</i>	VV	VV	V

Nunn assumes that these two tiers are sufficient:

In Nunn's notation, long vowels get the same notation as a pair of short vowels. The left-hand and right-hand examples above will thus get the same representation for the relevant vowel letters (Nunn 1998: 32-4), and there is no basis for the diaeresis rule to distinguish the two sets:

(30)	<i>orth-CV-tier</i>	<CVVC>	<CC VC VVC>
	<i>letter tier</i>	b a a l	b l a z o e n
		B a ä l	k a n o ë n

In a notation with three tiers, digraphs relate as an entity to the phonemic CV-tier, as represented above for <aa> and <ie>. From an orthographic point of view, these are two letters, but they correspond to one sound, which is what prevents application of the diaeresis rule.

Another argument that at least two tiers are needed comes from the representation of the glide /j/, which can be represented by a consonant letter or by a vowel letter. Representation of the glide by a vowel letter may result in an ambiguous letter string, later to be disambiguated by diaeresis placement, cf. *boeien* [bujən] and *Bedoeien* [beduwin]:

(31)	<i>orth-CV-tier</i>	<C V V V V C>	<C V C V V V V C>
	<i>letter tier</i>	b o e i e n 'cuffs'	B e d o e i e n 'Bedouin'
	<i>phon-CV-tier</i>	C V C V C	C V C V V C

The orthographical CV-tier offers no opportunity to distinguish both instances; it is the phonological tier that shows where a syllable boundary is present: CVCVC will be syllabified as CV.CVC, whereas CVCVVC will be syllabified as CV.CV.VC, with subsequent addition of the diaeresis.

In sum: a representation with an orthographical and a phonological CV-tier presents the opportunity to distinguish orthography and phonology in those cases that are not isomorphic. Graphotactic rules need both kinds of information. For instance: <ij> is one letter for a long vowel, <ch> is a combination of two letters for one consonant. Moreover, some phonological information is carried over in order to distinguish digraphs from pairs of monographs (<aa> for a long vowel or for two consecutive short ones) and to disambiguate <i> as either a vowel or a consonant. Two conclusions follow. First, the orthographic representation is a multi-tiered representation, and information from different levels is needed in the formulation of autonomous spelling rules. Second, the fact that phonology and orthography distinguish consonants and vowels in similar ways cannot be a coincidence. Rather than being completely autonomous, the orthography is partly a copy of the phonology.

6.2. Syllabification

From the Consistency Hypothesis it follows that orthographic rules should form an autonomous component; only one level can be relevant for a given writing system, and graphotactic rules should all refer to this level only. Later stages of the phonological derivation should not

be relevant to a writing system. Orthographic syllabification, however, is sometimes dependent on the spoken form.

Take a look first at instances of spoken and written syllabification in Dutch showing that different rules are involved. When speaking, especially in non-emphatic contexts, syllables may be formed from parts of different morphemes within a word or sometimes of different words in compounds, phrases, or sentences. In writing, syllable boundaries respect all word boundaries and some morpheme boundaries; cf. (32), in which dots or spaces indicate orthographic syllable boundaries and hyphens indicate phonological syllable boundaries:

(32)	<i>orthography</i>	<i>phonology</i>	
	ont.aarden	ɔn-tardən	lit. 'to de-earth', 'to degenerate'
	aard.appels	ar-dapəls	'potatoes'
	heel.al	he-ləl	lit. 'whole-all', 'universe'
	wel.is.waar	wɛ-lis-war	lit. 'it is true', 'indeed'
	'k ga 't 'm zeg.gen	kxɑ-təm-zɛ-χɛn	lit. 'I will-it-him tell', 'I will tell him that'
	'k wist 't	kwɪs-tət	'I knew it'

The examples (33) show that syllabification in phonology is based on the segmental string only, whereas syllabification in orthography refers also to morpheme boundaries. Because of this, the orthographic syllable boundaries need not coincide with the phonological ones in complex word forms, cf. the mismatches of phonological and orthographic syllables in the b-examples of (33):

(33)	<i>phonology</i>	<i>orthography</i>	
a.	[ra:r-stə]	raar.ste	stem <i>raar</i> , suffixes <i>-st</i> and <i>-e</i> 'weirdest'
b.	[sxa:r-stə]	schaars.te	stem <i>schaars</i> , suffix <i>-te</i> 'scarcity'
a.	[ko:r-tjə]	koor.tje	stem <i>koor</i> , suffix <i>-tje</i> 'small choir'
b.	[ko:r-tjə]	koord.je	stem <i>koord</i> , suffix <i>-je</i> 'small cord'
a.	[diploma-tjə]	diploma.tje	stem <i>diploma</i> , suffix <i>-tje</i> 'small certificate'
b.	[diploma-tjə]	diplomaat.je	stem <i>diplomaat</i> , suffix <i>-je</i> 'small diplomate'
a.	[dɔm-stə]	dom.ste	stem <i>dom</i> , suffixes <i>-st</i> and <i>-e</i> 'stupidest'
b.	[frɔn-stə]	frons.te	stem <i>frons</i> , suffix <i>-te</i> 'frowned'

In orthography, prefixes form a separate orthographic syllabification domain (cf. *ont.aarden* in (32)), as do consonant initial suffixes (*-te*, *-st*, *-tje*, *-je* in (33)). Another instance of a mismatch of phonological and orthographic syllables can be found in words with <i> representing a glide. Interestingly, words such as *baaien* [ba-jən] 'bays' and *bajes* [ba-jəs] 'prison', which are similar in all relevant phonological respects, get the different orthographic syllabification *baai.en* and *ba.jes*. The generalization is, that in orthography a syllable boundary is present after <i> that represents a glide, even if this would imply that more orthographic syllables are created than are present in speaking: *financieel* consists of four written syllables, but has only three phonological ones: *fi.nan.ci.eel* [fi-nan-sjel]. Similarly, after <u> representing a glide an orthographic syllable boundary is present, cf. *linguïst* with three written syllables, but only two phonological ones: *lin.gu.ist* [lɪŋ-wɪst].

The differences between syllables in writing and speaking may not lead to the conclusion that orthographic syllabification is based on the string of letters and morpheme boundaries alone. As with speaking, syllabification is dependent on the maximal onset principle, cf. *kar.tel* – *ka.trol*, *her.pes* – *ci.pres*, *hel.pen* – *di.ploma*. Clusters of consonants of decreasing sonority must be split; clusters of increasing sonority may form the onset of the following syllable. This can be captured by autonomous rules only if information about the sonority of the sounds is carried over to the letters. Alternatively, phonological syllabification may form the input to orthography. There is some evidence that the latter option is the correct one.

Observe that for some morphemes, variation in pronunciation occurs. For instance: *bio-* and *syn-* are pronounced [bi-jo], [bi-jɔ], [sin], or [sɪn]. In such cases, orthographic syllabification follows the most common pronunciation for that morpheme in words. That is, the prescribed orthographic forms *bio.sfeer*, *bios.coop*, *sy.no.niem*, *syn.er.ge.tisch* correlate with the common pronunciations [bi-jo-sfe:r], [bi-jɔs-kop], [si-no-nim] and [sɪn-ʔer-ʒe-tis]. The variation co-occurs with vowel length: when <o> is a long vowel, the syllable boundary follows immediately thereafter, when <o> is a short vowel, the syllable ends in <s>. Pronunciation and writing therefore go hand in hand.

If phonological syllabification is indeed the input of orthographic syllabification, then Nunn's autonomous spelling rules can no longer be considered autonomous. The context of orthographical syllabification must include information about phonological syllables. Her two-step analysis remains unchallenged, of course.

6.3. Degemination

In adjectives formed by the suffix *-s* following a stem ending in *-s*, degemination is systematically represented, but it is not represented in similar genitives, cf. the adjective *Parijs* and the genitive *Parijs'*, both with the morphological structure stem+s.

- (34) a. *Adjective forming -s*
 een Amsterdams huis <Amsterdam + s> 'an Amsterdam-like house'
 een Parijs bonbonnetje <Parijs + s> 'a Parisian bonbon'
- b. *Genitive forming -s*
 Amsterdams grachten <Amsterdam + s> 'Amsterdam's canals'
 Parijs' wegennet <Parijs + s> 'Paris's road system'

In phonology, the adjective forming *-s* and the genitive forming *-s* are nondistinct, as one would expect given a framework where phonology and morphosyntax form different components of the grammar, e.g. prosodic phonology (Selkirk 1984, Nespor and Vogel 1986). At the stage where degemination applies, information about morphological categories such as the distinction between the derivational adjective forming suffix *-s* and the inflectional genitival *-s* is no longer present. The spelling examples given above show that orthography distinguishes the two phonologically similar suffixes: degemination applies in both cases, but an apostrophe forms a trace of the deletion site in case the suffix is genitival.

Of course, it will be interesting to answer the question why the degemination rule is reflected in one class of words and not in the other. The answer can be found presumably in the inflection available for adjectives, cf.:

- (35) Amsterdamse huizen ‘Amsterdam-like houses’
 Parijse bonbonnetjes ‘Parisian bonbons’
 *Parijs’e bonbonnetjes

Adding inflectional *-e* changes the function of the apostrophe. Before or after a space, apostrophes are unambiguous traces of deletion, cf. (36a). But when an apostrophe is surrounded by letters, it usually functions as a linking element, cf. (36b):

- (36) a. *apostrophe next to space = trace of deletion*
 ’k ga naar huis < ik ga naar huis ‘I will go home’
 Als ’t regent < Als het regent ‘When it rains’
 Max’ regenjas < axs regenjas ‘Max’s raincoat’
- b. *apostrophe between letters = linking sign*
 baby’tje ‘small baby’
 menu’s ‘menus’
 NP’s ‘NPs’
 A4’tje ‘piece of A4-paper’

Examples of apostrophes as traces of deletion in between letters exist, cf. (37), but this use is highly restricted. It only occurs in idiosyncratic abbreviations (a frequently used name) or in contractions (indicating a schwa sound):

- (37) *apostrophe between letters = trace of deletion*
 A’dam < Amsterdam (not a general rule)
 ik heb d’r gezien < ik heb haar gezien ‘I saw her’ (general, but restricted to schwa)

This may explain why the apostrophe is not used in adjectives: the apostrophe in *Parijs’e bonbonnetjes* would be interpreted as a linking sign, not as a trace of deletion.

A consistent pattern could arise also when genitive *-s* would be treated as the adjectival *-s*, without an apostrophe indicating the deletion site. However, this would cause ambiguity in names for which a stem with and without an *-s* occurs, cf. *Philips boek* (from: *Philip+s*) and *Philips’ boek* (from: *Philips+s*). When both requirements must be met (readability for apostrophes and an unambiguous representation for proper names), there is no way out but to represent degemination inconsistently.

Here again, alternative approaches may rescue the Consistency Hypothesis. One may propose that there are two different rules of *s*-degemination, one for derivation (the adjective forming *-s*) and one for inflection (the genitive *-s*), or that degemination applies cyclically, such that there will be a stage in the phonological derivation at which adjectival degemination has applied and genitive degemination has not. But such an approach would deny that the true explanation of the inconsistency lies in an output constraint on the use of apostrophes and in the need to represent names unambiguously. The idea of an Orthographically Relevant Level that forms for each language the pivot between the spoken and written mode is interesting because it restricts the options otherwise available, but provides a framework in which explanations will not be found.

6.4. Stress

Stress in words is generally not reflected in Dutch orthography. There is one exception: the spelling of morpheme-final /i/. In simplex forms, this /i/ is spelled <ie> in both stressed and unstressed syllables. However, when a suffix is added that begins with a vowel, <ie> is spelled <i> in unstressed position and remains <ie> in stressed position, cf. the following examples (underlines indicate stress):

- (38) olie, menie 'oil, minium' oliën, meniën 'to oil, to minium'
drie, strategie 'three, strategy' drieën, strategieën 'threes, strategies'

Stress assignment is based on syllables, cf. (39a), which not always coincide with morpheme boundaries, cf. (39b):

- (39) a. artistiek – artisticiteit 'artistic, artisticity'
winkel – winkelier 'shop, shopkeeper'
vijand – vijandig 'enemy, hostile'
leraar – lerares 'teacher, mistress'
neger – negerin 'Negro, Negress'
afwas – afwasbaar 'washing-up, washable'
- b. art+ist+ic+iteit ar-tis-ti-ci-teit
winkel+ier win-ke-lier
vijand+ig vij-an-dig
ler+ar+es le-ra-res
neger+in ne-ge-rin

Hence, information of two kinds of structure are needed in orthography: morpheme structure and syllable structure. Of course, one may call the combination of both a consistent level of linguistic representation, but then a wider interpretation of this notion is intended.

7. The Orthographically Relevant Level in Dutch

In the above sections on autonomous spelling rules, three kinds of arguments were forwarded against a single Orthographically Relevant Level for Dutch. First, phonological and morphological information is carried over to a later stage where autonomous spelling rules apply. Second, phonological information of shallower levels is needed for the proper application of the autonomous rules. Finally, the autonomous rules are near copies of the phonological rules. This leads to the conclusion that phonological rules and orthographical rules are closely related, as if the orthographical component is working in parallel with the phonological component.

Dutch orthography challenges the claim that languages universally obey the Consistency Hypothesis. Whereas the conversion to the orthography is most profitably described when the level of phonemic representations forms the input, one may not be led to the conclusion that only the information of this single linguistic level provides sufficient information. As the sur-

vey of section 2 shows, morphology is also relevant for Dutch; i.e., the domains are provided by morphology, whereas the conversion rules refer to phonemes.

For some morphemes, a more superficial phonological level forms the input to the spelling representation. As arguments for the different status of such morphemes cannot be found (except that they are spelled differently), such spellings form counter-evidence to the Consistency Hypothesis. Furthermore, the spelling of native and non-native morphemes is derived by slightly different rules for both sets. Since non-native complex forms are spelled more as if they are simplex forms, a different morphological level seems to be relevant for these two strata of Dutch – a deeper level for the spelling of native formations (i.e. separate morphemes) and a more superficial level for the spelling of non-native words (i.e. complex forms).

For the autonomous spelling rules, both phonological and orthographical information is relevant. The phonological information relates in some cases to quite superficial levels, for instance, at which syllable structure and stress are present. The conclusion must be that such rules have letters as their target, but that they refer to the phonological context, also from later stages than the phonemic level. The rules are not as autonomous as suggested by their name, witnessed by the fact that orthographic rules look like phonological rules. When punctuation rules are considered to belong to the class of autonomous rules (and nothing seems to contradict this), then global information from morphology, syntax, and semantics also is needed in the second stage of deriving the surface orthographical representation. The Orthographically Relevant Level thus contains information of nearly all components of the grammar.

Given an incremental approach to language processing (see, for instance, Levelt 1989 and Levelt et al. 1999) all information of earlier stages is retained in more superficial stages. Then, the Consistency Hypothesis would be met trivially by the phonetic representation in all languages, since information about phonemes, morpheme boundaries, morphological categories, syntactic categories, and the like would be present at that level. No restrictions would follow from the Consistency Hypothesis, which then becomes superfluous.

The Consistency Hypothesis of course would be a natural constraint on the interface of linguistic components. Given the above argument, however, the conclusion must be that the orthographic component is not related in such a natural way to the components of the spoken mode in Dutch, which represents information on syllables, morphemes, phonemes, phones, and even parts of syntax and semantics. Perhaps this explains why the acquisition of literacy is so different from the acquisition of natural languages, including sign languages.²

² I am grateful to Geert Booij, Mirjam Ernestus, Anneke Nunn, Richard Wiese, and my both co-editors for their comments on an earlier version of this paper.

References

- Asher, Ronald E. & J.M.Y. Simpson (1994): *The Encyclopedia of Language and Linguistics*. Volume 9. Oxford, New York, Seoul, Tokyo: Pergamon Press.
- Beurden, Lisan van (1987): Playing Level with Dutch Morphology. In: Frans Beukema & Peter Coopmans (eds.): *Linguistics in the Netherlands 1987*. Dordrecht: Foris, 21-30.
- Booij, Geert E. (1987): The Reflection of Linguistic Structure in Dutch Spelling. In: Philipp A. Luelsdorff (ed.) *Orthography and Phonology*. Amsterdam, Philadelphia: Benjamins, 215-225.
- Booij, Geert E., Camiel Hamans, Gerard Verhoeven, Frida Balk & Ch.H. van Minnen (1979): *Spelling*. Groningen: Wolters-Noordhoff (= *Spektator Cahiers* 2).
- Boyes Braem, Penny (1995): *Einführung in die Gebärdensprache und ihre Erforschung*. Hamburg: Signum.
- Bredel, Ursula (this volume): The Dash in German.
- Carney, Edward (1994): *A Survey of English Spelling*. London, New York: Routledge.
- Daniels, Peter T. & William Bright (1996): *The World's Writing Systems*. New York, Oxford: Oxford University Press.
- Fabb, Nigel (1988): English Suffixation is Constrained only by Selectional Restrictions. *Natural Language and Linguistic Theory* 6: 527-539.
- Giegerich, Heinz J. (1999): *Lexical Strata in English. Morphological Causes, Phonological Effects*. Cambridge: Cambridge University Press.
- Haas, William (1976): Writing: the Basic Options. In: William Haas (ed.): *Writing without Letters*. Manchester: Manchester University Press, 1-28.
- Haas, Wim & Mieke Trommelen (1993): *Morfologisch handboek van het Nederlands*. 's Gravenhage: SDU Uitgeverij.
- Haeringen, Coenraad B. van (1937/1962): 'Spelling pronunciations' in het Nederlands. *De Nieuwe Taalgids* 31, reprinted in: *Neerlandica, verspreide opstellen*. Den Haag: Damen, 37-60.
- Hay, Jennifer (1999): *Causes and Consequences of Word Structure*. Ph.D. dissertation North Western University.
- Heuven, Vincent J.J.P. van (1978): *Spelling en lezen. Hoe tragisch zijn de werkwoordsvormen?* Assen, Amsterdam: Van Gorcum.
- (this volume): Effects of Diaeresis on Visual Word Recognition in Dutch.
- Heuven, Vincent J.J.P. van, Anneke H. Neijt & Maarten Hijzelendoorn (1994): Automatische indeling van Nederlandse woorden op basis van etymologische filters. In: *Spektator* 23: 279-291.
- Jespersen, Otto (1909): *A Modern English Grammar. Part I: Sounds and Spellings*. Copenhagen: Munksgaard.
- Levelt, Willem J.M. (1989): *Speaking*. Cambridge and London: MIT press.
- Levelt, Willem J.M., Ardi Roelofs & Antje S. Meyer (1999): A Theory of Lexical Access in Speech Production. *Brain and Brain Sciences* 22: 1-75.
- Maas, Utz (2000): *Orthographie – Materialien zu einem erklärenden Handbuch zur Rechtschreibung des Deutschen*. Ms. University of Osnabrück.
- Meisenburg, Trudel (1996): *Romanische Schriftsysteme im Vergleich*. Tübingen: Narr.
- Neef, Martin (1996): *Worddesign*. Tübingen: Stauffenburg.
- (this volume): The Reader's View: Sharpening in German.
- Nespor, Marina & Irene Vogel (1986): *Prosodic Phonology*. Dordrecht: Foris.
- Nunn, Anneke M. (1998): *Dutch Orthography. A Systematic Investigation of the Spelling of Dutch Words*. The Hague: Holland Academic Graphics.
- Plag, Ingo (1999): *Morphological Productivity. Structural Constraints in English Derivation*. Berlin/New York: Mouton de Gruyter.
- Rietveld, Antonius C.M. (1983): *Syllaben, klemtonen en de automatische detectie van beklemdoende syllaben in het Nederlands*. Ph.D. dissertation, Nijmegen University.
- Sampson, Geoffrey (1985): *Writing Systems: A Linguistic Introduction*. London: Hutchinson.

- Sandra, Dominiek, Steven Gillis & Georges De Schutter (1996): The Interaction of Phonological and Morphological Knowledge in Children's Intuitive Syllabifications. In: Mieke Beers & Sandra Peters (eds.): *Phonological Development: Different Perspectives*. (= *Antwerp Papers in Linguistics* 96), 17-31.
- Schreuder, Rob, Anneke Neijt, Femke van der Weide & R. Harald Baayen (1998): Regular Plurals in Dutch Compounds: Linking Graphemes or Morphemes? *Language and Cognitive Processes* 13: 551-573.
- Seidenberg, Mark S. & Michael K. Tanenhaus (1979): Orthographic Effects on Rhyme Monitoring. *Journal of Experimental Psychology: Human Learning and Memory* 5: 546-554.
- Selkirk, Elisabeth O. (1984): *Phonology and Syntax. The Relation between Sound and Structure*. Cambridge, Mass.: MIT Press.
- Sgall, Petr (1987): Towards a Theory of Phonemic Orthography. In: Philipp A. Luelsdorff (ed.): *Orthography and Phonology*. Amsterdam, Philadelphia: Benjamins, 1-30.
- Sproat, Richard (2000): *A Computational Theory of Writing Systems*. Cambridge: Cambridge University Press (*Studies in Natural Language Processing*).
- Wells, John Christopher (1982, 1995): *Accents of English, Volume I: An Introduction*. Cambridge, New York, Melbourne: Cambridge University Press.
- Wilbur, Ronnie (1987): *American Sign Language: Linguistic and Applied Dimensions*. Boston: Little, Brown and Co.
- Winkel, Lambert A. te (1863): *Ontwerp der Spelling*. Leiden: D. Noothoven van Goor.
- Zonneveld, Wim (1983): Lexical and Phonological Properties of Dutch Voicing Assimilation. In: Marcel van den Broecke, Vincent van Heuven & Wim Zonneveld (eds.): *Sound Structures. Studies for Antonie Cohen*. Dordrecht, Cinnaminson: Foris, 297-312.