Chapter 5
ASPECTS OF DUTCH ORTHOGRAPHY AND READING

Vincent J. van Heuven

LETTER TO PHONEME CONVERSION

When trying to place orthographies along a scale, ranging from e.g., Finnish, as an illustration of a near-optimally regular set of letter-phoneme correspondence rules, to English, which has traditionally been considered as highly irregular, Dutch would presumably be located toward the regular end of this scale.

Complete regularity is usually taken in the sense of perfect one-to-one letter-phoneme correspondences. However regular these correspondences in Dutch may be, there is no single one-to-one letter-phoneme correspondence in Dutch orthography, contrary to what has been suggested in the literature (for an opposing view see Damsteegt, 1976).

Relevant Letter Symbols in Dutch Orthography

Dutch spelling has been changed by governmental decree a number of times this century, and, in fact, public clamor for additional changes is still heard today. The changes involved are all intended to alleviate the problems that confront the (would be) writer of Dutch. One important category of spelling problems is constituted by the spelling of nonnative words, notably words of Latin and French origin. I would like to dispense with this class of words in this survey for a number of reasons:

The spelling of Latinate words will in practice be mainly a matter of rote learning

Latinate words are nearly always polysyllabic and of low frequency, and are therefore not dealt with in the initial reading stages

There is a tendency to use alternative, native spellings, which although not always officially approved by the government, are socially acceptable.

A consequence of leaving out Latinate words is that 23 letters are used in Dutch orthography, the usual set of 26 minus *q*, *x*, and *y*. However, *c* remains only as part of the digraph *ch*.

Phoneme Inventory of Dutch

Phonemic analyses of Dutch vary to some extent as to the number of relevant contrasts. There are some major issues in Dutch phonology that may affect the outcome of a phoneme count.

The Analysis of Trigraph Vowels (Sometimes Referred to as Long Diphthongs) The issue at stake is whether such complex vowels as [a:i] or [e:u] in *vlaai* [vla:i] 'pastry' or *geeuw* [χe:u] 'yawn' are to be thought of as single phonemes or as sequences of two phonemes. Most analyses today take the position that these complexes should be regarded as a long vowel followed by a separate consonant /j/ or /w/ (for a survey and references in English see Brink, 1970), which is the practice followed in this chapter.

Loan Words from the French Supposedly due to French occupation of the Netherlands in the early nineteenth century, a large number of French loan words have been incorporated in the Dutch lexicon, retaining more or less their original pronunciation. Although there is considerable overlap of Dutch and French near-equivalents in the respective phoneme inventories, a number of foreign sounds had to be added to the Dutch phonemic system. Moreover, these loan words have also retained much of their original French spelling, although partial adaptation to the standard conventions of Dutch orthography has been allowed, or even advocated, in the course of the last two decades. In spite of this, the Dutch spelling of French loan words is often considered uneducated, especially in the eyes of the older and middle generations.

Since there is still no official uniform settlement on the spelling of French loan words, part of the French letter-phoneme correspondence rules are duplicated in the Dutch system. This, of course, causes serious problems to anyone who wants to read and write Dutch.

Even though many of these loan words are frequently used in Dutch, and may have no native equivalent, e.g., the words *douche* [du:ʃ] 'shower', or *serre* [sɛ:rə] 'glass house', I omit these words, too, from further consideration, mainly on the strength of the argument that their spelling is, understandably, not dealt with in the initial spelling and reading stages in Dutch primary schools.

The Phonological Status of the Neutral Vowel Schwa [ə] There is still some dispute on the correct interpretation of a central vowel in Dutch which can only occur in unaccented syllables: [ə]. It has a variety of representations in spelling: *e, ij, i* (for details cf p. 64). In terms of vowel quality it does not differ from another central vowel [œ], spelled *u,* with which it is allegedly in opposition. Error analyses of spontaneous spellings of preschool children (van Rijnsoever, 1977) indicate that the distinction is highly obscure to untrained speakers of the language. Certain radical spelling reform proposals, which have not (yet) been accepted, advocate reduction to one phoneme, and consequently to one symbol.

Table 1. Conversion from single vowel symbols to phonemes

Symbol	Phonemic value	
	Initial/medial position	Final position
a	/ɑ/ *al* 'already', *pak* 'grab'	/a:/ *pa* 'dad'
e	/ɛ/ *er* 'there', *weg* 'road'	/ə/ *de* 'the'
i	/ɪ/ *in* 'in', *kin* 'chin'	
o	/ɔ/ *os* 'ox', *bos* 'wood'	/o:/ *vlo* 'flee'
u	/œ/ *uk* 'tod', *buk* 'stoop'	/y:/ *nu* 'now'

Velar Fricatives Perceptually there is no difference between voiced and voiceless velar fricatives (van Heuven & van den Broecke, 1977), although the most recent textbooks on Dutch phonetics (Nooteboom & Cohen, 1976) maintain that the distinction exists. From a transformational point of view there are compelling arguments to accept the difference as an underlying distinction, which is captured in the spelling in a rigorous fashion. Analysis of spelling errors of first graders, however, shows that the distinction is lost: usually only the symbol for the voiced variant, *g,* is written, although it is the voiceless sound that is pronounced throughout. The reason for this spelling behavior is probably that the symbol for the voiceless fricative, *ch,* a digraph, is taught at a later stage than the voiced symbol.

Incidentally, the voiced/voiceless distinction, which has always been neutralized in final position in all Germanic languages except English, and which has hardly any functional load in Dutch, seems to be disappearing in all fricatives in initial positions, and even medial positions as well. In fact, in certain urban dialects the distinction is lost completely, which leads to additional spelling difficulties for such dialect speakers.

Letter-Phoneme Correspondences in Dutch

In view of the facts that have been given in the preceding sections, it seems reasonable to limit the discussion of Dutch letter-phoneme correspondences to monosyllabic, native words, which, as a matter of fact, also constitute the subject matter for the first grade reading and writing skills at primary schools. An extension to polysyllabic words follows in the next section.

Under the restrictions made here, pronunciation of Dutch words is fully predictable from spelling, although the converse does not hold. Using 23 letters to cover 33 phonemes (34 including /ə/), among which are 15 (16) vowels that have to be represented by only 5 vowel symbols: *a, e, i, o* and *u,* some conventions had to be used in order to have combinations of vowels represent the various vowel phonemes.

Strictly speaking, the 18 consonant letters could quite adequately serve to represent the 18 consonant phonemes, but this turns out not to be the case.

Table 1 shows how to relate single vowel symbols to phonemes, in syllable initial and medial position on the one hand, and in final position on the other.

Table 2. Conversion from digraph vowel symbols to phonemes

Symbol	Phoneme	Position		
		Initial	Medial	Final
aa	/a:/	*aal* 'eel'	*zaak* 'case'	
au	/ɑu/	*au* 'ouch'	*paus* 'pope'	*au* 'ouch'
ee	/e:/	*eed* 'oath'	*leed* 'sorrow'	*zee* 'sea'
ei	/ɛi/	*eik* 'oak'	*reis* 'voyage'	*kei* 'boulder'
eu	/ø:/		*reus* 'giant'	*reu* 'dog'
ie	/i:/	*iep* 'elm'	*biet* 'beet'	*wie* 'who'
ij	/ɛ:/	*ijs* 'ice'	*pijn* 'pain'	*rij* 'row'
oe	/u:/	*oer* 'prehistoric'	*boer* 'farmer'	*koe* 'cow'
oo	/o:/	*oog* 'eye'	*boog* 'bow'	
ou	/ɑu/	*oud* 'old'	*goud* 'gold'	*kou* 'cold'
ui	/ʌy/	*uil* 'owl'	*kuil* 'pit'	*bui* 'shower'
uu	/y:/	*uur* 'hour'	*buur* 'neighbour'	

The reason why a single *a, e, o, u* (and in foreign words, *i*) could be used to represent long vowels in final position, is that no short vowels are permitted in that position by phonotactic constraint. Notice further that long /e:/ is not expressed by *e* in final position, as this represents the neutral vowel /ə/. As a final remark on Table 1 it should be noted that *u* in nonfinal position represents long /y:/ if followed by *w*. All other vowel representations are digraphs (see Table 2).

Curiously enough, *au* in final position only occurs in onomatopeic *au* [ɑu] 'outch'; in all other words it is followed by *w*. The digraph *ou,* which represents the same phoneme, may or may not be followed by *w* in final positions, leading to homophones: *jou/jouw* [jɑʊ], 'you/your' and *kou/kouw* [kɑʊ], 'cold/chew'.

The letter-phoneme conversion for consonants is relatively simple. Unless stated otherwise, any consonant symbol has the value given to it in the International Phonetic Alphabet (IPA). There are, however, some exceptions:

ch is pronounced as [χ], *ng* (only occurring in final position) is pronounced as [ŋ], and *n* before *k* is pronounced as [ŋ].

A final consonant symbol representing a voiced obstruent is pronounced as as a voiceless sound: *word* [wɔrt] 'become' and *heb* [hɛp] 'have'.

The symbol traditionally representing a voiced velar stop, *g,* is pronounced as a voiceless velar fricative, irrespective of its position in the word: *graag* [χra:χ] 'voluntarily'.

The final combination *dt,* which occurs in verb forms only, is pronounced as [t]: *wordt* [wɔrt] 'becomes'.

i and *u* as the last element of a trigraph vowel (cf. p. 58) represent the consonants [j] and [w] respectively. However, *u* is not used to designate a semivowel after another symbol *u.* This graphotactic constraint was

apparently introduced to avoid such illegal letter sequences as **ruu*[1] [ry:w], which is correctly spelled *ruw* 'rough'. [1]Notice that the symbol *u* is not geminated here, as should be expected, since it is no longer in syllable final position.

In a comprehensive treatment of the letter-phoneme correspondences in the Dutch monosyllable, there is at least one further complication that should be dealt with. There are quite a few words in Dutch that are spelled with only one vocalic center, but are regularly pronounced as two syllables. This phenomenon arises when a liquid or nasal is followed by a nonhomorganic consonant in the postvocalic position. In such cases an epenthetic unstressed neutral vowel [ə] is inserted between the two consonants. This rule explains, for example, why the name of the bulb flower *tulp* ['tœləp] was understood as 'tulip' by English speakers. Thus *melk* 'milk' is pronounced ['mɛlək], and *berg* 'mountain' as [bɛrəχ], but *lamp* 'lamp' remains [lɑmp]. There is a small group of historically motivated exceptions to the epenthesis rule that need not concern us here; for a survey of some relevant data and extensive references see Brink (1970, pp. 145–146).

In conclusion to this section, one allophonic variation, viz., the influence of /r/ on preceding [+tense] vowels must be mentioned. For high vowels the effect is mainly a matter of lengthening. For mid vowels, which are clearly diphthongized in all other environments, the change in quality disappears. After /a:/ the change is usually towards [ɛ] ('t Hart, 1969). Young children, who tend to listen to sounds very analytically, are often found to spell **vir* for *veer* 'feather' or **bur* for *beur* 'lift', which contain the vowel symbols that normally represent the short vowels with the same quality.

Underlying Principles of Dutch Orthography

The Phonemic Principle Unless another rule takes precedence, all contrastive sounds are uniquely represented by letters or letter combinations. Thus, audibly quite different sounds such as long vowels followed by final /r/ have the same spelling as when occurring in other environments, because they are allophonic variants. The sounds /ʌy/ and /ɑu/, which speakers of English find very difficult to distinguish, are spelled differently because they function contrastively in Dutch.

The Congruence Principle Stem morphemes are not spelled differently when they undergo predictable phonetic changes in morphological alternants. The final devoicing rule for obstruents is ignored in the spelling using this principle; similarly, a number of low level phonetic assimilation processes are not reflected in the spelling.

[1]An asterisk (*) denotes a nonoccurring form.

The Analogy Principle Some deep level suffixes are consistently represented in the orthography, even if they are absent in pronunciation, for example, the practice of spelling *dt* at the end of verbs. The inaudible addition of *t* signals, for example, the grammatical function of third person singular indicative present, on the analogy of the audible addition of /t/ to other verb stems such as *zwemt* [zwɛmt] '(he) swims'. As a consequence of this principle, the first and third persons of verbs with stems on *d* are homophonous: (ik) bid [bɪt] '(I) pray' versus *(hij) bidt* [bɪt] '(he) prays'. Notice, however, that no inaudible addition of *t* is allowed to stems ending in *t* themselves: *(hij) schiet* [sχ:t] '(he) shoots', but not **schiett*.

The Etymological Principle Different letter codings with identical sound values may be employed to capture distinctions that existed in older stages in the development of the language. The spelling alternatives for /ɛi/ *(ei, ij)* and /ɑu/ *(au, ou)* are examples of this principle. Contrary to English orthography, where a similar principle may be used (cf. Chomsky & Halle, 1968), these distinctions seem to have no significance in a synchronic description of Dutch (morpho-) phonology. This renders the etymological principle an unnecessary source of error in the eyes of many specialists on Dutch orthography, and reduction to single representations of these phonemes has often been proposed.

On the basis of the particulars given so far it turns out that automatic letter-phoneme conversion for native Dutch monosyllabic words is a relatively simple task, and will always have error-free results. In order to determine the phonemic value of a particular letter, a moving context window should span three adjacent letters. This would at least partly explain why Dutch is more regular than English. In English many letter-phoneme correspondences are idiosyncratic: *great* (/ei/) versus *beat* (/':/), or even worse, *read* with either /':/ or /e/, in which case a decision cannot be made without recourse to higher order morpho-syntactic information. The phonemic value of Finnish letters, however, can probably be derived by taking only two adjacent letters into account, as would also be the case in Spanish orthography.

Polysyllabic Words

The first important complication of letter-phoneme conversion in polysyllables is the following convention, pertaining to the spelling of certain vowels and all single letter consonants. As we have seen in the discussion of monosyllables, a vowel which is normally spelled with reduplication *(aa, ee, oo, uu)* is simplified (reduced to one letter) in syllable final position, or, as it is traditionally called, in an 'open' syllable. One of the few clear principles of Dutch syllabification is that, unless a word boundary intervenes, a single intervocalic consonant is the first element of the next syllable. This implies that when, for example, *beek* 'brook' is pluralized (by adding *en* [ən]), the

syllable division is *bee-ken,* leaving *ee* in final position. Consequently the stem vowel is now spelled with single *e: beken.* To complicate matters further, vowel simplification does not take place when the next syllable begins with a digraph consonant: *goochel* [χo:-χəl] 'juggle'.

Conversely, a single letter consonant is reduplicated in the intervocalic position after a stressed short vowel. Thus the plural of *gek* 'fool' is spelled *gekken* and *sok* 'sock' as *sokken.* Understandably, this rule is a direct consequence of the vowel simplification rule. If the intervocalic consonant would not be reduplicated, the stem vowel would be assigned the value of a long vowel.

This consonant/vowel reduplication system is not dealt with until the second grade, and spontaneous spellings produced when the rules have not yet been taught, reveal no tendencies to use such conventions. The exact advantages of this system, if any, are neither known, nor have they ever been investigated. One advantage would be that average word length in Dutch is now slightly lower than when a Finnish kind of spelling would be used (always reduplicate letters representing long phonemes). In an average Dutch text, vowel simplification occurs about twice as often as consonant reduplication, with relevant cases in roughly every fifth word.

In spite of this, the system clearly goes against the grain of the morphophonemic basis of the Dutch writing system: it obscures a direct one-to-one letter-phoneme correspondence and it interferes with uniform spellings of both free and bound morphemes. There is, however, no clear indication that these conventions constitute significant spelling or reading problems, once the rules are dealt with.

This practice of geminating vowel and consonant symbols, ultimately, warrants my earlier claim that there is no single one-to-one letter-phoneme correspondence in Dutch.

Unsoluble Cases

Correct letter-phoneme conversion in Dutch polysyllables presupposes two extra sources of information:

Where are the syllable boundaries?
Which syllables are unstressed?

Unfortunately, proper syllabification and stress assignment is not always possible on the basis of formal criteria. Phonological syllabification is dependent on internal word boundaries. A difference in position of an internal boundary alone may create a homographic pair:

boordraad	*boor-draad*	[bo:rdra:t]	'spiral thread of a drill'
	boord-raad	[bo:rtra:t]	'council on board a ship'
haardrek	*haar-drek*	[ha:rdrɛk]	'filthy mess of hairs'
	haard-rek	[ha:rtrɛk]	'fire guard'

Because the *d* in these pairs is alternatively in syllable final or in initial position, the final devoicing rule may or may not apply, creating phonemic differences. Obviously, the correct placement of internal word boundaries in these compounds can only be done on the basis of the semantic context in which these words occur.

As another example of unavoidable error, consider the following compound: *darmonderzoek* 'intestinal examination'. Assuming that the typical Dutch syllable is a CVC sequence, the straightforward syllabification of this word would be *dar-mon-der-zoek,* which would be pronounced [dɑrmɔndərzu:k]. However, the word is a compound of *darm* 'bowel' and *onderzoek* 'examination', so that an internal word boundary occurs between *m* and *o*. Since *rm* is now a final cluster, the epenthesis rule becomes applicable, giving the pronunciation [dɑrəmɔndərzu:k]. In cases like these each attempt at syllabification should be preceded by some lexical decision process in order to ensure correct division.

Brandt Corstius (1970) has evaluated his syllabification program, which is currently used by Dutch newspapers for automatic typesetting, on the basis of the error rate obtained in the automatic syllabification of a 43,712 word token (4114 types) corpus of newspaper texts (van Berkel, Brandt Corstius, Mokken, & van Wijngaarden, 1965). He found that 64 of the 4,114 types (1.6%) were hyphenated incorrectly. It seems to me that these results are somewhat misleading, as a large part of the words in the corpus were monosyllabic. Brandt Corstius mentions a percentage of 54, which (probably) refers to word tokens.

A second source of ambiguity is the decision whether certain syllables in a word are part of the stem or separate morphemes. If they are separate morphemes, they will be unstressed. Normally, stress differences have no phonemic consequences, but there are some exceptions:

ij, which is normally pronounced as [ɛi], must be pronounced as [ə] in the suffix *lijk* [lək], which has a variety of uses. Unfortunately, there is also a stem form *lijk* [lɛik] 'corps', so that a correct analysis of a word like *kinderlijken* depends on semantics: when *ij* is unstressed, [kɪndərləkən], the meaning is "childlike persons'; when stressed, [kɪndərlɛikən], it means 'children's corpses'.

e stands for [ə] when unstressed, but represents [e:] in stressed, open, non-word final syllables:

geren	*'geren*	[χe:rən]	'taper(ing) out'
	ge'ren	[χərɛn]	'repeated act of running'
bedelen	*'bedelen*	[be:dələn]	'beg'
	be'delen	[bədə:lən]	'hand out'

i, which regularly represents [ɪ], is pronounced as [ə] in the suffix *ig,* as in *prachtig* [prɑχtəχ] 'beautiful'.

Kok (1972a, 1972b) describes a program for automatic letter-phoneme conversion for Dutch texts, and mentions an error rate of 6%, on the basis of word types in a 44,299 word token corpus (9,380 types). Nine percent of the errors were due to incorrect syllabification, 16% to the foreign status of certain words, 24% to the absence of diacritics that are normally written in Dutch, and 51% to incorrect stress assignment. These figures illustrate once more that Dutch orthography is problematic as soon as one goes beyond the scope of the monosyllable.

PROBLEMS IN THE INITIAL READING STAGES

There are two sources of literature that may contain information on the aspects of Dutch orthography that constitute learning problems in the initial reading stage: 1) experimental studies on initial reading, carried out by educational psychologists and 2) textbooks for prospective primary school teachers.

In this discussion of possible reading problems the survey is limited to technical or mechanical reading, the process of correctly pronouncing sequences of letters (words), no matter whether or not the reader understands what their meaning is.

The experimental studies referred to, however, limit their scope even further to sound pure words (Kooreman, 1974, 1975), words selected from an artificially constrained lexicon in which only one-to-one phoneme grapheme correspondences hold. Although a number of interesting learning problems can be tackled in this way, the procedure eliminates any possibility to investigate reading problems specifically due to the peculiarities of Dutch orthography. In other words, studies on the basis of such a constrained orthography might as well have been conducted in another language with a writing system with a one-to-one grapheme-phoneme relationship. These studies are therefore not taken into further consideration.

Research specifically directed toward investigating whether certain types of letter-phoneme conversions are intrinsically more difficult to learn than others has never been carried out in the field of Dutch orthography. In order to get some idea as to what might be the outcome of such an investigation, one could perhaps analyze initial readers' performances on a number of mechanical reading skill tests. It should be pointed out, however, that none of the tests was constructed with this goal in mind. One might, as a matter of fact, wonder whether the authors of such tests were at all aware that quite a few items in their tests contained deviations from one-to-one letter-phoneme correspondences, and potentially constituted additional difficulties, over and above those inherent to the mechanical reading skill as such.

Unfortunately I have not been able to perform any such post hoc analyses, partly because no score distributions of the tests are readily available in the literature and partly for reasons of time.

For the second source of information we may turn to the more or less intuitive claims as to what causes reading problems in Dutch, that can sometimes be found in textbooks for prospective teachers. Although most textbooks (and teacher's handbooks that come with reading primers) contain some thoughts on this matter, the most comprehensive listing of reading difficulties is given by Caesar (1971, p. 36):

Reading vowels followed by *-r,* which leads to a change in quality
Reading digraph consonant symbols: *ch* and *ng*
Reading words in which a final voiced symbol is pronounced as a voiceless sound: *bed, heb*
Reading words with nongeminate digraph vowels (but not with geminate digraphs)
Reading the final sound in trigraph vowels (*aai, eeu,* etc)
Reading [ə] for *e* in final position
Reading words in which a single letter vowel represents a long vowel phoneme in word final position and open syllables in general
Pronouncing [ə] for *e* or *ij* in unstressed syllables
Reading words with syllable final nonhomorganic [+sonorant] [+consonant] clusters (epenthesis rule)
Reading final *nk* as in *bank* 'bank, bench', which a beginning reader would be tempted to pronounce as [bɑnək] by erroneous application of the epenthesis rule, and consequently fail to recognize
Reading polysyllabic words
Reading words deviating from the simple CVC structure, where CVCC is claimed to be easier than CCVC

Although this is a rather elaborate list, one might wonder if it is exhaustive. It is the result of an a priori crude analysis of letter-sound discrepancies and the impression of classroom errors that were informally observed. A more rigorous analysis of letter-sound discrepancies will undoubtedly reveal additional potential difficulties; in fact, a number of omissions will be detected if one compares the listing here with the discussion of letter-phoneme correspondences given earlier.

All but a few of the listed problems are traced to complications in letter-phoneme correspondence. One problem has to do with difficulties that come up when longer and more complex words are read, and two more problems originate in markedness of syllable structure. I should like to enlarge on this last issue, and discuss an experiment in which deviations in syllable structure were methodically investigated as a source of problems.

Rispens (1974), who investigated the relationship between children's abilities to synthesize words from isolated sounds, and reading perfor-

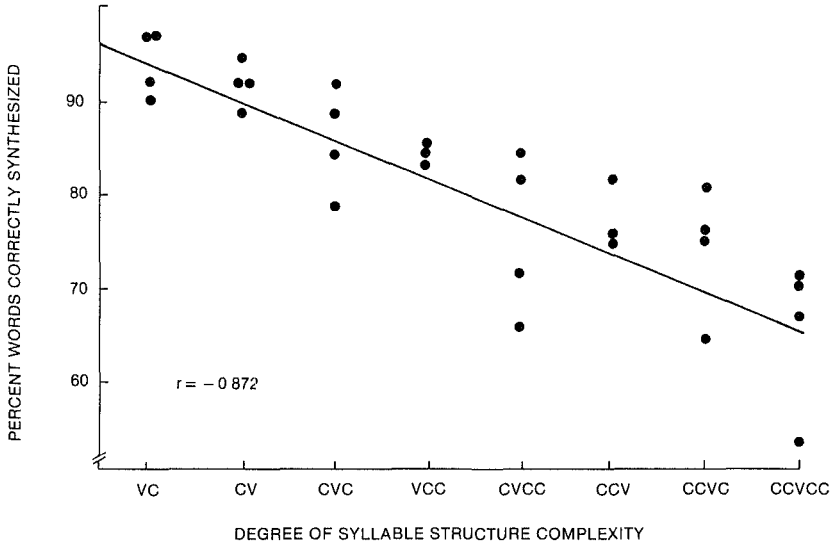

Figure 1. Percentage of correctly synthesized (blended) words as a function of syllable structure complexity. Each dot represents the results for one test word, based on the responses of 91 subjects (adapted from Rispens, 1974).

mance, hypothesized the following order of difficulty in monosyllabic words: 1) VC, 2) CV, 3) CVC, 4) VCC, 5) CVCC, 6) CCV, 7) CCVC, and 8) CCVCC. His crucial test measured difficulty of synthesis (blending individual sounds to a word) as a function of syllable structure complexity. A group of 91 initial readers responded to 32 test words presented as sequences of isolated sounds, 4 different test words per syllable structure type. The results are as indicated in Figure 1, in which the percentage of correctly synthesized words is plotted along the vertical axis, and syllable structure complexity increases along the horizontal axis.

It appears from the graph that there is a fairly strong relationship between the two variables ($r = 0.875$). However, Rispens' data are confounded by word frequency. Correlation of percent correctly synthesized words with log transformed word frequency, as established in a recent 600,000 word count of Dutch texts (uit den Boogaart, 1975), turns out to be 0.524. If my claim that high frequency words are easier to synthesize than low frequency words (and, in fact, correlation of syllable complexity and log word frequency was -0.559) is accepted, then Rispens' findings, suggestive though they may be, should be regarded with caution. A rerun of his experiment with adequate control for word frequency seems to be in order.

Insufficient experimental data are available to warrant any definite conclusions as to what constitutes a reading difficulty in Dutch. Nevertheless, the author feels that initial reading instruction would greatly benefit

from clearer insights both in the consequences of discrepancies of letter-phoneme correspondences and in the effects of syllable complexity.

EXPERIMENTS ON THE EFFECTS OF SOME CHARACTERISTICS OF DUTCH SPELLING ON SENTENCE PROCESSING IN ADULT READING

Dutch orthography is basically a morphophonemic writing system, in which uniform spelling of stems and affixes is preserved, even if this leads to complications in letter-phoneme correspondence. As a result, there are a large number of morphologically complex homophones whose spelling differences are caused by differences in underlying morphological structure. Correct spelling of such words is only possible if the writer realizes that alternative morphological analyses may correspond to a particular sound sequence, selects the right alternative, and reflects this in his spelling.

An example of this is the inaudible spelling difference between first and third persons of verbs in the present tense with stems ending in *d* (cf. pages 57–65). It turns out that the correct spelling of verb forms, which is taught from third grade onward, presents tremendous problems, which situation has once been called "the tragedy of the verb forms" (Van der Velde, 1956). All spelling reform proposals that are currently entertained suggest simplifications on these points.

In a discussion of spelling reforms two distinct groups of users should be kept in mind: those who have to learn to read and write, and those who have mastered the system. It is at least conceivable that certain characteristics of an orthography that are difficult to learn, may greatly facilitate the processing of text in a later stage (see Frith & Frith, chapter 18, this volume; Lukatela & Turvey, chapter 15, this volume).

In 1974–1977 I worked on a project that aimed to investigate if reflection of underlying morphological structure in homophonous verb forms might be beneficial to adult readers. Because the correct morphological analysis of the verb forms (and consequently their spelling) depends on the syntactic structure of the sentence in which they occur, it was hypothesized that the resulting redundancy of verb inflection and syntactic structure might guide the reader when processing a sentence.

From a theoretical point of view, Dutch orthography is an interesting system to conduct this kind of experiment in, because it provides the means to compare different ways in which underlying morphological distinctions may be reflected in spelling and pronunciation:

type I: An underlying distinction is maintained in the spelling and in pronunciation: *(ik) speel* [spe:l] '(I) play'/*(hij) speelt* [spe:lt] '(he) plays'.

type II: An underlying distinction is neutralized in pronunciation but maintained in the spelling, thus creating a homophone: *(ik) bid* [bɪt] '(I) pray'/*(hij) bidt* [bɪt] '(he) prays'.

type III: An underlying opposition is lost both in spelling and in pronunciation: *(ik) schiet* [sχi:t] '(I) shoot'/*(hij) schiet* [sχi:t] '(hij) shoots'.

The experiment to be discussed presently is one of series, carried within the framework of the project referred to. It has been described in more detail together with a number of related experiments (van Heuven, 1978). Preliminary reports of a more informal nature have appeared in the Progress Reports of the Institute of Phonetics of Utrecht University (van Heuven, 1976, 1977a, 1977b).

Cue Value of Tense Marking Suffixes in Plural Finites in Silent Reading

In Dutch plural finites, grammatical person is not expressed, but there is a formal difference between present and past tense, which may be reflected in pronunciation and spelling in each of the three ways mentioned. Weak verbs add *ten* or *den* to the verb stem in the past if the stem ends in an underlying voiceless or voiced segment respectively. The present tense plural is formed by suffixing *en* to the stem. Thus the three opposition types are as follows:

type I: *werken* [wɛrkən] 'work'/*werkten* [wɛrktən] 'worked'
zwaaien [zwa:jən] 'wave'/*zwaaiden* [zwa:jdən] 'waved'
type II: *feesten* [fe:stən] 'have a party'/*feestten* [fe:stən] 'had a party'
branden [brɑndən] 'burn'/*brandden* [brɑndən] 'burned'
type III: *dutten* [dœtən] 'nap'/ditto 'napped'
wedden [wɛdən] 'wager'/ditto 'wagered'

Note that the gemination of *t* and *d* in the past tense of type II is caused by the analogy principle (the past tense morpheme is spelled uniformly in type I and type II). In type III, however, gemination is caused by the general spelling convention stating that stressed short vowels must be followed by two consonant symbols (cf. pp. 57–65).

The actual stimulus material consisted of 40 sentences constructed by systematic variation of opposition type (see above) and three other variables, to be discussed presently.

Sentences were complex, beginning with a temporal clause followed by the main clause. The subclause started with a conjunction that can be used with both present and past tense. Both clauses contained a finite. On account of concord the tense of the second clause is predictable from the tense of the first. The temporal clause contained the crucial verb, the second clause contained a strong verb, in which the difference between the tenses was visible only in one short letter in the middle of the word. Assuming that the perception of such strong verbs takes the same amount of time in the present and in the past (e.g., *trekken/trokken* 'draw/drew'), the speed and accuracy with which readers may decide that the tense of the second clause is compatible with that of the first, will depend on the relative salience of the tense cue contained in the crucial verb form.

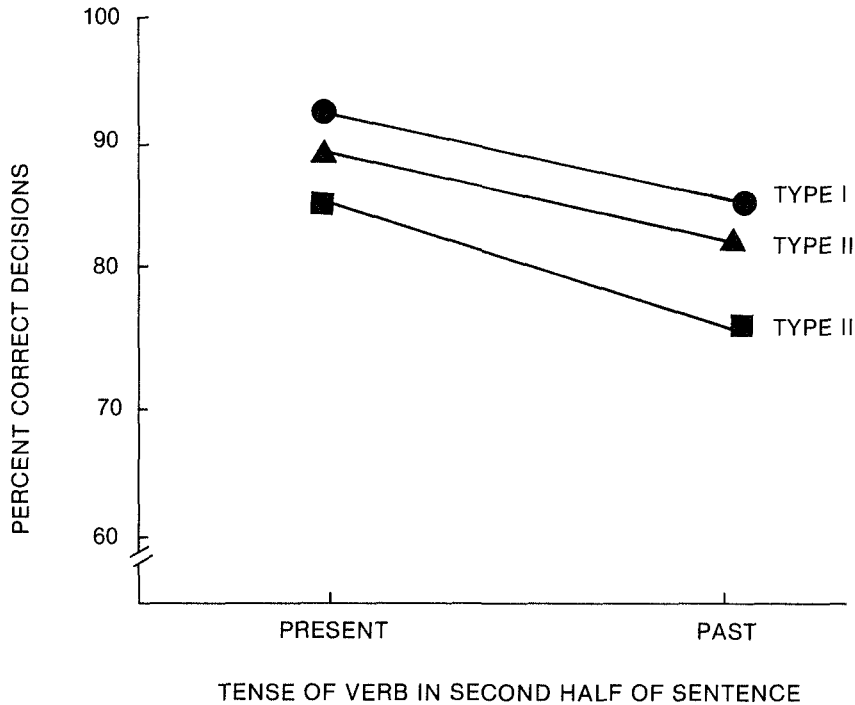

Figure 2. Percentage of correct decisions as a function of opposition type (type I: audible and visual difference; type II: only visual difference; type III: no difference) and tense of verb in second half of the stimulus sentence. Each mark is based on 624 responses (312 for type III).

Four different sentence frames were used, two of which contained verbs with *t*-stems (underlying voiceless stem final segment) and two more with *d*-stems (with voiced stem final segment). With the exception of the type III crucial verbs, whose tense cannot be varied, both first and second verbs could be either present or past, leading to violation of tense concord in half of the type I and type II sentences.

The sentences were presented twice, in different quasi-random orders, to a group of 39 readers (male and female staff and students at Utrecht University, right-handed and with normal vision, paid volunteers, and all native speakers of Dutch) with the aid of a line stepper (Bouma & De Voogd, 1974). The subclause was presented for 1 sec, and then replaced by the main clause, which remained visible until the subject responded by pressing one of two buttons marked "right" or "wrong." The subjects were instructed to decide for each pair of clauses, as fast and as accurately as possible, whether or not the combination was a correct sentence. One-half of the subjects pressed right for "right" and left for "wrong"; positions were reversed for the other half.

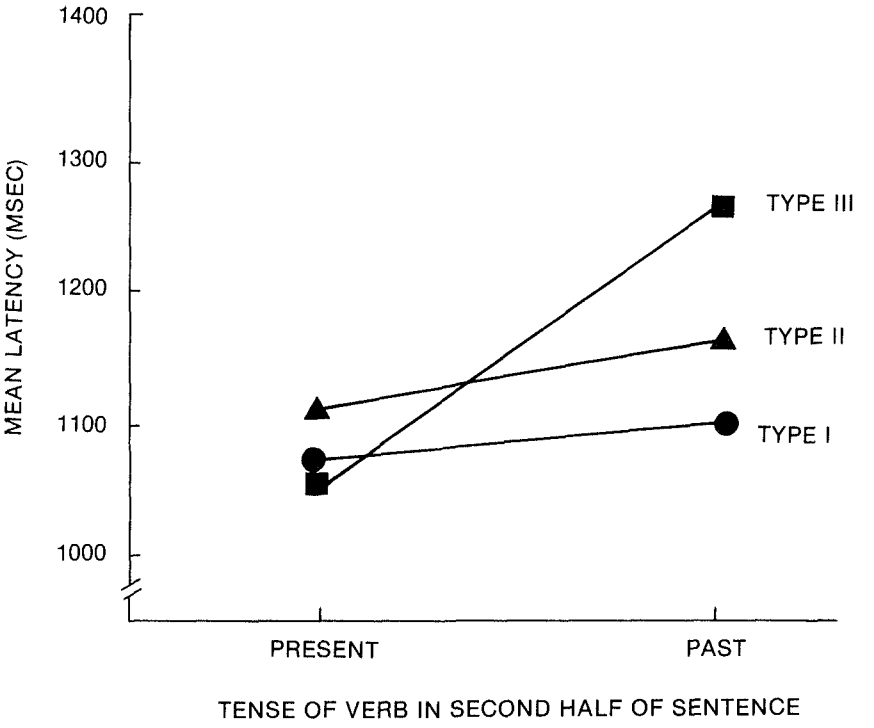

Figure 3. Decision latency (msec) for correct responses only, as a function of opposition type and tense of verb in second half of stimulus sentence.

Response latency (time interval between moment of presentation of the second part of a stimulus and moment of response) was measured, and the correctness of the response was established. Since type III crucial verbs are compatible with both present and past tense continuations, a "wrong" response here was regarded as an error. The results are given in Figure 2 for mean percentage correct decisions and in Figure 3 for mean latency obtained with correct decisions only.

The general level of accuracy is 87% correct. Type I verbs are responded to more accurately (90% correct) than type II (87% correct), which in turn is superior to type III (82% correct). There is a small effect due to the four different sentence frames used, but differences disappear completely when the sentence frames are lumped together two by two on the basis of *t*- and *d*-stems. Present tense continuations are responded to with 90% accuracy, past tense continuations with only 84%.

The particular sentence frame has no influence at all on latency, nor is there an effect when the sentences are combined on *t*- and *d*-stem characteristics. Shortest latencies are obtained with type I oppositions (1086 msec),

longest for type III (1165 msec), and intermediate latencies are found for type II (1121 msec). Present tense continuations lead to faster correct responses than past tense (1080 msec versus 1157 msec). However, the advantage of the present tense continuation is much stronger for type III contrasts than for the other types.

DISCUSSION OF RESULTS

In general it appears that the percentage of correct decisions and their response latencies supply more or less parallel information, which is what one expects to be the case. The results of this experiment clearly supported our earlier findings that the grammatical cue provided by a potentially audible contrast (i.e., type I) is more effective than when the difference exists only in the spelling (type II). However, the results forced us to qualify an earlier claim that type II cues would not be any different from the control type containing no cue at all (type III). Rather it should be said that the inaudibility of a spelled difference reduces its cue value to some extent.

This finding seems to be in line with later weak versions of the speech recoding hypothesis as formulated by, for example, Kleiman (1975) and Baddeley (1976, 1979). In their view speech recoding is not strictly necessary for reading, but it may provide a means to perform tasks in which memory plays a role. There is a striking resemblance between one of Kleiman's experiments, from which it appeared that subjects had great problems in judging the semantic acceptability of strings of consecutively presented words while recoding to speech was precluded by imposing a concurrent shadowing task, and our experiment. In our experiment the tense information derivable from the verb in the first clause had to be remembered for some length of time to be able to arrive at a decision after viewing the second part of the sentence.

Past tense forms in type II oppositions *(feestten, brandden)* were responded to more accurately and faster than type III forms *(wedden, dutten)*. This means that the gemination of the medial consonant letter is perceived in most cases, and that it is used as a cue to determine the grammatical tense of the clause. However, when forms such as *dutten* and *wedden* were presented, i.e., without informative suffixes and hence compatible with both present and past tense, correct decisions that a present tense continuation of the sentence is grammatical, are taken faster than with a past tense continuation.

In view of these facts one must assume that the perception of the word final letter strings *tten* or *dden* is not automatically associated with past tense. Rather it appears as if the subject performs a more sophisticated analysis of the verb forms: only if *tt* or *dd* is preceded by another consonant

symbol, does he know that gemination is due to the past tense suffix. If the letter preceding the *tt* or *dd* is a single vowel symbol, he realizes that gemination is the direct consequence of a general spelling rule, and that no information relevant to performing his task is given.