

ORTHOGRAPHY AND PHONOLOGY

ORTHOGRAPHY AND PHONOLOGY

Edited by

PHILIP A. LUELSDORFF
Universität Regensburg

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM/PHILADELPHIA

1987

Library of Congress Cataloging in Publication Data

Orthography and phonology.

Papers presented at a workshop held at the Fifth International Phonology Meeting, held in Eisenstadt, Austria, June 25-28, 1984.

Bibliography: p.

1. Language and languages -- Orthography and spelling -- Congresses. 2. Grammar, Comparative and general -- Phonology -- Congresses. I. Luelsdorff, Philip.

P240.2.075 1987 411 87-9361

ISBN 90 272 2039 5 (alk. paper)

© Copyright 1987 - John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

CONTENTS

List of Contributors	vii
Preface	ix
Towards a Theory of Phonemic Orthography	1
<i>Petr Sgall</i>	
The Description of Spelling-to-Sound Relationships in English, French, and Russian: Progress, Problems, and Prospects	31
<i>Bruce L. Derwing, Tom M.S. Priestly, and Bernard L. Rochet</i>	
The Automatic Phonetic Transcription of English Text	53
<i>S.G.C. Lawrence, B.J. Williams, and G. Kaye</i>	
Phonological Access to Data Bases	65
<i>L. Hitzenberger</i>	
On Linguistic Error	77
<i>Philip A. Luelsdorff</i>	
A Preliminary Report on a Pilot Investigation of Greenlandic School Children's Spelling Errors	101
<i>Birgitte Jacobsen</i>	
The Orthographic Representation of Nasal Vowels in Acehnese	131
<i>Mark Durie</i>	
Phonological Recoding in the Reading Process	151
<i>Hartmut Günther</i>	
Orthography as a Variable in Psycholinguistic Experiments	171
<i>Bruce L. Derwing and Maureen L. Dow</i>	
Spelling and Dialect	187
<i>Gunnel Melchers</i>	
The Reflection of Linguistic Structure in Dutch Spelling	215
<i>Geert E. Booij</i>	
Zur alphabetischen Orthographie als Gegenstand der Sprachwissen- schaft	225
<i>Herbert Penzl</i>	

LIST OF CONTRIBUTORS

- Geert E. Booij, vakgroep Algemene Taalwetenschap, Vrije Universiteit, P.O. Box 7161, 1007 MC Amsterdam, The Netherlands
- Bruce L. Derwing, Department of Linguistics, Faculty of Science, The University of Alberta, Edmonton, Alberta, Canada T6G 2H1
- Maureen L. Dow, Department of Linguistics, Faculty of Science, The University of Alberta, Edmonton, Alberta, Canada T6G 2H1
- Mark Durie, Department of Russian and Language Studies, The University of Melbourne, Parkville, Victoria 3052, Australia
- Hartmut Günther, Max Planck Institut für Psycholinguistik, Wundtlaan 1, NL - 6525 XC Nijmegen, The Netherlands
- Ludwig Hitzenberger, Linguistische Informationswissenschaft, Universität Regensburg, Postfach 397, 8400 Regensburg 1, Federal Republic of Germany
- Birgitte Jacobsen, Institute of Linguistics and Institute of Eskimology, University of Copenhagen, Copenhagen, Denmark
- G. Kaye, M.P. 137, Speech Systems, IBM UK Scientific Centre, Athelsten House, St Clement Street, Winchester SO23 9DR, United Kingdom
- S.G.C. Lawrence, M.P. 137, Speech Systems, IBM UK Scientific Centre, Athelsten House, St Clement Street, Winchester SO23 9DR, United Kingdom
- Philip A. Luelsdorff, Institut für Anglistik, Universität Regensburg, Postfach 397, 8400 Regensburg 1, Federal Republic of Germany
- Gunnel Melchers, Engelska Institutionen, Universitet Stockholms, Fack, 104 05 Stockholm 50, Sweden
- Herbert Penzl, Queen Beatrix Chair of Dutch Language, Literature and Culture, Department of German, University of California at Berkeley, Berkeley, California 94720, U.S.A.

Tom M.S. Priestly, Department of Slavic and East European Studies, The University of Alberta, Edmonton, Alberta, Canada T6G 2H1

Bernard L. Rochet, Department of Romance Languages, The University of Alberta, Edmonton, Alberta, Canada T6G 2H1

Petr Sgall, Matematicko-fyzikální fakulta, Katedra Applikované Matematiky, Malostranské nám. 25, 11 800 Praha 1, Czechoslovakia

Briony J. Williams, M.P. 137, Speech Systems, IBM UK Scientific Centre, Athelsten House, St. Clement Street, Winchester SO23 9DR, United Kingdom

PREFACE

This volume contains the full versions of the papers delivered and discussed during the workshop on Orthography and Phonology at the Fifth International Phonology Meeting held in Eisenstadt, Austria, from June 25-28, 1984. We are grateful to the workshop participants for their contributions and stimulating discussion and to The Organizing Committee of the Meeting, in particular Wolfgang U. Dressler, Oskar E. Pfeiffer, and John R. Rennison, for their professional commitment and efficient organization.

Collected here are several papers devoted to various aspects of the orthography/phonology interface. Topics covered include spelling-to-sound correspondence for English, French, and Russian, the design of a generative phonology for orthography data-base access, the linguistic sign and orthographic and phonological error, the analysis of Greenlandic school children's spelling errors, the orthographic representation of phonemic nasalization and its implications for prosodic theory, the psycholinguistics of phonological recoding in reading, orthography as a variable in psycholinguistic experiments, spelling and dialect, orthography and the typology of phonological rules, and orthography and historical phonology.

Petr Sgall formulates basic definitions required by a theoretical description of grapheme systems grounded in the phonemic principle. These include 'grapheme', 'protographeme', 'subgrapheme', 'complex grapheme', 'epigrapheme', and 'alphabet'. Two conditions are seen to underlie every phonemic orthography: the uniqueness of pronunciation of a grapheme and the uniqueness of spelling of phoneme. Departures from such biuniqueness lead to definitions of the additional concepts of 'bigrapheme', 'trigrapheme', and 'quasibigrapheme'. The central ingredient of a rule of orthography as well as a rule of pronunciation is considered to be the relation Orth, defined as a pair of strings, the leftmost string consisting of phonemes, the rightmost one of graphemes. Deviations from phoneme-grapheme biuniqueness are seen to lie on two dimensions, one of complex-

ness (from letter to sound) and one of univocality (from sound to letter). Sgall makes interesting use of these notions drafting (the, to the best of our knowledge, first) scales of complexity and univocality. These can be used as a basis for typologizing alphabetic orthographic rules within and across languages, studies in language acquisition, and discussions of orthographic reform.

The paper by Derwing, Priestly, and Rochet summarizes the author's research on the systematic description of "reading rules", i.e. spelling-to-sound correspondences, in Russian, English, and French. Two types of reading rules are posited: (1.) "spelling rules" converting the standard orthography into a regularized orthography in which redundancies and inconsistencies are minimized and (2.) "pronunciation rules" relating the regularized orthography to a "phoneticized orthography". General problems under discussion include: rules and exceptions, rules and lists, and the simplicity vs. practicality of linguistic descriptions.

Lawrence, Williams, and Kaye describe an automated text-to-speech transcription system for English. Their rules convert standard orthographic representations into phonemic, allophonic, and phonetic transcriptions. Rule ordering is crucial to both this system and the one described above; the essential difference lies in the additional set of orthographic regularization rules in the reading rules.

In response to the need for access to written name files, where the correct spelling of the name might not be known, Hitzenberger sketches the design of a generative phonology for data-base access. The general idea is to translate an orthographic representation of a name into sequences of phonological feature bundles and then translate these feature sequences back into all possible associated orthographic representations. The orthographic access possibilities are increased by reducing the number of features, hence the number of phonemes.

Luelsdorff presents an extended view of the notion "linguistic error" within the Bühlerian organon-model of language, a discussion of the notion "possible linguistic error", a critique of a current notion of linguistic sign and a revision thereof designed to predict the major occurring linguistic error types, including those orthographic and phonological.

Jacobsen's paper is a pilot study of Greenlandic school children's spelling errors in the new Greenlandic phonemic orthography. She focuses on quantity errors in the spelling of vowels and consonants, errors resulting from the orthographic overdifferentiation of uvularized and non-uvularized vowels, insertions and omissions of ⟨r⟩, and the correlation between quantity-error frequency, word length, and word complexity.

Durie discusses the interesting case of the orthographic representation of nasal vowels in Acehnese, an Austronesian language spoken in the north of Sumatra. Although nasal vowels are phonemic in Acehnese, their phonemic status is obscured in the orthographies which have thusfar been proposed. Durie concludes that the Acehnese speaker's representation of distinctive nasality is prosodic, not segmental, with the general implication that segmental phonetic constraints may be encoded non-segmentally.

Günther argues against a psychological interpretation of the view that writing represents sound and sound represents meaning. Several models of the reading process are reviewed, phonological recoding and the internal lexicon defined, indirect access to the internal lexicon discussed, and several lexical decision experiments, the pseudoword argument, and recent conclusions from neurolinguistics critiqued.

The contribution by Derwing and Dow shows orthography to be a significant experimental variable in aspects of psycholinguistic research. Some examples of orthographic effects are reviewed and the results of two experiments presented. In the first, students required to provide written rhymes for words visually presented were found to be influenced by the spelling of the test word in their choice of rhymes. In the second, students required to record the number of speech sounds in token words exhibited no presentation-mode effect. The authors conclude that orthography can be an important influence on phonological judgments depending upon the type of task.

In an essay on spelling and dialect, Melchers discusses eyedialect and the case of Shetland English and the feasibility of dialect spelling in general. Under consideration are dialect writing as a field of study, the limitations and possibilities of orthography to represent dialect, Shetland spelling conventions, the use of dialect in the teaching of reading, and the crossdialectal adequacy of conventional English spelling.

Booij shows that an orthography may represent different levels of language, not just speech. He proceeds by establishing a typology of phonological rules for Dutch and determining which kinds of rules the Dutch spelling system reflects. Dutch spelling is shown to represent underlying forms, phonetic forms, and intermediate levels of phonological representation and to prefer a hyphen after a phonological-word boundary to a hyphen after a syllable boundary.

Finally, Penzl, focusing on orthography as the main evidence for historical phonology, deals with questions of orthography in the history of German.

TOWARDS A THEORY OF PHONEMIC ORTHOGRAPHY

Petr Sgall

0.1 The objective of the present study is to formulate the basic definitions that seem to be necessary for a theoretical description of graphemic systems based on the phonemic principle. Although exceptions and deviations can be found perhaps in all the writing systems of European languages, in most cases it is possible to base their descriptions on the relationship between the grapheme and the phoneme. Not only in the relatively regular systems, such as e.g. those of Spanish or of Polish, but also in those biased by phonetic writing (e.g. Croatian, White Russian, cf. fn. 21) or by relations between graphemes and morphemes (or syllables), as in English or French, it seems most useful to describe the exceptions or deviations as such, the phonemic principle serving as the basis of the description. We want to show in the next paragraphs that this holds true even if the autonomy of the level of graphemics is acknowledged; without this acknowledgment a systematic description of the mentioned exceptions (or of the rules concerning the use of punctuation marks, underlining, capitals, spaces, etc.) would not be easy.

We are not interested in syllabic or pictographical, hieroglyphic and similar writing systems, and do not pay attention here to such transition systems as those of the Devanagari alphabet, of Ivrit or of Arabic graphemics. Also the difference between writing and print is neglected in the present study, so that *written* often stands here also for *printed*. It also has to be noted that it is not the written norm as such, which is in the centre of our attention (see esp. Vachek, 1939; 1972, on the specific functional status of the written norm); we are interested here, first of all, in the graphemic level itself (i.e. in one of the ingredients of written language) and in the relationships between this level and the corresponding levels of the spoken language.

We attempt a characterization of the grapheme level as a level in its

own right; however, as we shall see, it is not possible to characterize this level without paying due regard to phonemics, and also to other levels. Without this we would not be able to distinguish between characters or graphemes, on the one hand, and between punctuation marks, inverted commas, etc. on the other; we would not even have the possibility of distinguishing phonemic writing from syllabically or morphemically based writing systems. Thus, distinguishing between prototypes or, in classical Praguian terminology, the centre, core and periphery, we can see the prototypical (central) relationship in the correspondence between graphemes and phonemes, other cases (where graphemic means correspond to phenomena from other levels) being described as peripheral.

0.2 A comprehensive survey of the development of the investigations of graphemics and its relationships to spoken language is presented by Ruszkiewicz (1976), who evaluates the different approaches from Eastern Europe as well as from the West. He duly states that the first systematic formulations of the theory of graphemics were those by Baudouin de Courtenay. This great Polish linguist was also the first to formulate — as early as in 1871 — the following principles of orthography:

- (a) phonetic, reflecting the pronunciation of words,
- (b) etymological, reflecting the origin of words, i.e. their morphemic structure (in modern terminology),
- (c) historical, referring “to those cases in which the spelling of the word cannot be justified by the two remaining principles but by tradition only” (see Ruszkiewicz 1976: 27).

The most important turning point of the investigations of graphemics that is connected with European structural linguistics is based on the recognition that writing is an autonomous system, relatively independent of the spoken norm, see Vachek (1933; 1939; 1959), who quotes Artymovyc and at a certain point also Frinta as his predecessors, and Ruszkiewicz (1976: 16), where also Bradley’s relevant formulations from the beginning of this century are reported. Vachek has systematically elaborated the view according to which the written norm of a language, although historically secondary (see esp. Vachek, 1948), is relatively independent of the spoken norm of the given language.

Without going into the details of the development of different views on graphemics, let us just briefly summarize the survey by Ruszkiewicz, who states (pp. 37 to 46; 57) that the two main approaches are:

- (i) the relational one, according to which the grapheme is viewed as a

merely relational unit, “standing for” (expressing) a phoneme (Balinskaja, Bierwisch, Istrin, Gleason, Francis and others represent different standpoints within this trend);

(ii) the autonomistic approach, recognizing the grapheme as a unit of the written system (or norm) of language, which symbolizes mental experience directly, rather than via spoken language, and is based on its own oppositions (Vachek, 1949), on the basis of which the notion of grapheme can be defined; Ruszkiewicz prefers approach (ii), quoting also Leontjev, McIntosh, McLaughlin, Motsch, Penttilä, Stetson, Pulgram, Spang-Hanssen and Hammarström as adhering to similar views.

Ruszkiewicz then defines the grapheme as a class of non-contrastive allographs, where an allograph is a class of similar concrete written characters (p.47).

However, several remarks seem to be necessary here. We do not dwell on the fact that the use of the term “class” (although supported by Pulgram’s 1951 argumentation) does not seem to be fully appropriate, since it is first of all the type-token relation that is relevant; other points may be considered more important:

First, the notion of non-contrastiveness probably does not suffice as the main or single criterion for grouping characters into graphemes, i.e. the oppositions within graphemes do not constitute a sufficient basis for distinguishing the graphemes.¹ As Ruszkiewicz himself recalls (p. 39), quoting Gleason, the complementary distribution (which, in this case, instantiates the absence of contrast) of Greek -σ- and -ς is not sufficient for the assignment of these two characters to a single grapheme, since there are also other characters in the Greek alphabet that do not occur at the end of word forms, i.e. -ς is noncontrastive with more characters than just with σ.

Second, the question of capital letters is open to further discussion, since e.g. the relationship between *t* and *T* within a written norm certainly cannot be classed along with that between *t* and *p*. If we want to be able to use such formulations as “the grapheme *t* in the Roman alphabet”, taking into account that such two inscriptions as *LINGUISTICS* and *Linguistics* are equivalent on a certain level of the written norm, then it seems that the notion of a grapheme — since Baudouin understood as the basic unit of graphemics — should be preferably defined in such a way as to cover the capital as well as the small letter to which the capital corresponds.² From this it follows that the functional difference between a capital and the corresponding small letter (belonging to the core of the linguistic system in

marking proper names and the beginnings of sentences, and, on the other hand, having a stylistic character in other positions) should not be regarded as relevant for the definition of the grapheme, although other kinds of functional relevance should be (see the preceding paragraph about distinguishing -s from characters belonging to other graphemes).

In any case, the notion of grapheme itself as well as the individual graphemes can only be characterized with regard to the pronunciation-spelling relationship (cf. Bierwisch, 1971).

It is true that the historical primacy of spoken over written norms is neither absolute, nor directly relevant; in many cases, in ontological development, speakers come into acquaintance with the written shapes of morphemes before they know their phonemic representation (see Marcel, 1980: 402), and the influence between the two norms proceeds in both directions.³ Thus, writing is not only autonomous in that it does not reflect phonemic structures, but it can even be said, from a certain standpoint, to simply coexist with phonemics in its own right (see e.g. Uldall 1944: 14).

However, the following three points substantiate the view that rather than symbolizing mental experience directly, the written norm necessarily stands in a structural correspondence to the spoken one and is marked in its relationship to the latter — a view held also by Vachek (1939: pp. 413, 415; 1972: in press), who as one of the first understood the written norm as autonomous:⁴

(i) In the prototypical cases a grapheme corresponds to a phoneme; their relationship need not be univocal, but still, it appears as intuitive and empirically adequate to assume that the graphemic shapes of many morphemes are derived from the same underlying (phonemic) representations that underlie also their phonetic shapes, cf. Helberg (1974: 37ff), Weigl and Bierwisch (1970: 8f) and Nauclér (1983); we are aware that different approaches to phonemics can be chosen. The relationships between phonemes and graphemes are neither simple nor are they a complete basis for the identification of the graphemes, but they must be taken into account for the description of the graphemic shapes of such morphemes as E. *cut*, *mist*, *fine*, *lie*, *pre-*, *-ing*, *-istic*, and so on, not to speak about languages with a more “regular” orthography than English. The deviations, however numerous and varied they are, belong to the periphery rather than to the core of a system of phonemic orthography.⁵

(ii) The marked character of writing versus speaking is clearly reflected by the fact that writing systems are relatively artificial (Bazell:

1956); “while speech is acquired without formal teaching, reading and writing are skills to be taught” (Naucré, 1983, 595). The conventions underlying written norms (including not only graphemes, but also punctuation, abbreviations, etc.) are therefore accessible to conscious interventions to a degree to which rules of grammar (and phonemics) are not.⁶

(iii) Even in the cases when the written form does not correspond to phonemics in such a way as in the prototypical cases illustrated in (i), it corresponds to the structuring of another level of language — of morphophonemics (as is the case of E. *-s*, *-ed*),⁷ or morphemics (e.g. with *right* — *write*, as in lexical morphemes), or of syntax and the level of meaning (esp. with punctuation).⁸ There are not two systems of a language (a spoken and a written one) used by the same linguistic community, but rather a single system having one level of meaning, one of (surface) syntax, one of morphemics, one of phonemics (another of morphophonemics), and two parallel “output” levels: phonetics and graphemics. The parallelism is not complete from a structural point of view (see Bazell, 1956), and is complicated — in respect of its functional aspects — by the cases where (strings of) graphemes are derived directly from another level than from that of phonemics;⁹ cf. Fig. 1, where the lines denote the relationships of means (*signifiés*) to functions (*signifiants*), understood technically as “being directly derived from” (the dotted lines denote peripheral cases of this relation). The prototypical parallelism of the two levels is crucial for the alphabetical systems of writing to be distinguished from others (e.g. hierarchical ones).¹⁰

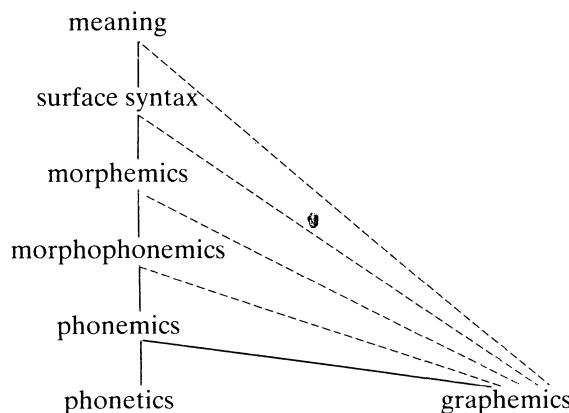


Fig. 1.

Therefore, even if the autonomy of graphemics as a level in its own right is acknowledged, it is not possible to define the basic notions of graphemics without taking into account their relationships to (a) the phonemes and (b) the units of other levels of the system of a language.¹¹ Otherwise one would be able only to describe characters, the similarities in their shapes (cf. Stetson's (1937) 'distinguishing traits') and in their distribution, or perhaps the relationships between typological sets of letters, handwriting, etc. This layer of "graphographemes" (see Volockaja et al., 1964; Heller, 1980: 94f) can perhaps be characterized as another level, the relationship of which to graphemics (or graphophonemics) is similar to that between phonetics and phonemics.

It should also be kept in mind that without paying due regard to the relationship between writing and phonemics it would be difficult even to distinguish between such basic kinds of graphemic systems as phonemic, syllabic or ideographic ones. As is well known, the fact that phonemic writing is so widely spread has to do with the number of phonemes being much smaller than that of morphemes or syllables; it has been noted that a graphemic system could also be based on distinctive features and that the drawbacks of syllabic writing are not absolute (Stuart, 1958).

In any case, the definitions of the basic units of graphemic system have to cope with such differences as those between simple characters and those containing diacritics, between small and capital letters, between standard types, italics, boldface types, etc., and also with underlining, punctuation marks, the blank, with the specific cases as individual graphemes added to an otherwise international alphabet (e.g. German ß), or the French accents not being used with capitals, and so on. Thus, it should not be assumed that a simple definition of a single notion of grapheme could suffice.

0.3 When discussing the relationships between graphemes and phonemes, we work here first of all with the classical notion of phoneme, as known from the Prague School. This notion seems to be more appropriate for our purposes than the abstract phoneme of generative phonemics (see Vachek, 1964, where the difference between the two approaches is discussed).¹²

If we are not mistaken, the classical phoneme can also be defined in terms of generative phonemics, although it does not constitute one of the main units of that approach. The European alphabets, derived more or less directly from the Ancient Greek one, seem to consist first of all of symbols reflecting phonemes in this sense, rather than morphophonemes (or

the abstract phoneme of generative phonemics). Even after Chomsky and Halle (1968) have shown that certain differences between phonemics and graphemics can be understood as much smaller if the abstract phoneme is used in the descriptions, it seems more typical that many quite regular phonemic alternations in different languages are rendered by pairs or graphemes (e.g. in Czech *matka* - *matce* the graphemes *k* and *c* reflect the two shapes of a single morphophoneme). Certainly, deviations can be found also in this respect, but regularly only a case of neutralization is reflected by a single grapheme in the European orthographies (German *d* renders also the phoneme /t/ in *Grad*, etc.)

After having presented a brief survey of some points of the background, we formulate in the next section a tentative, informal characterization of several notions that could underlie a theory of graphemics.

1.1 Before formulating definitions of such notions as alphabet, orthography, and spelling, we need to characterize a few more elementary notions. Since we are interested in phonemic systems, rather than in hieroglyphic or other orthographies, it seems appropriate to take as a starting point the fact that a written text has the shape of a string of slots some of which are occupied by characters (possibly accompanied by diacritics or by underlining), others by other symbols, still others being empty.

We understand here the concept of a slot in the string as a primitive notion; other primitive notions are: a symbol, which can occupy a slot (or a part of it, in some specific cases), and, furthermore, those used as basic notions in phonemics and other subdivisions of linguistics: a phoneme and its allophones or variants,¹³ a word-form, a text, a stereotype.

We can now characterize the following concepts:

A blank is a slot that is empty, not occupied by any symbol.

A character is a symbol the tokens of which are used in writing in such a way that (a) every token of the symbol occupies a slot, and (b) a token of the symbol can be used as a part of a string representing a word-form.¹⁴

In the prototypical cases a character is pronounced, i.e. the users of a natural language understand it as representing (by itself, or as a part of a string of characters) a phoneme (or a string of phonemes, see below): thus, when occurring in a text that is being read, the English grapheme *c* is pronounced as (a variant of) the phoneme /s/ in such word-

forms as *centre*, *reception*, as /k/ in *cook*, and together with a following *h* it is pronounced as /č/ in *check*, as /š/ in *machine*; such symbols as the comma or the hyphen are not pronounced in English, i.e. they are not immediately connected with phonemes (here we do not go into such questions as whether a pause can be viewed as a pronunciation of a comma). In peripheral cases a character is not pronounced (Fr. *h*, or older Russian Ъ), or pronunciation is absent in certain of its occurrences determined by the morphemic context, as e.g. E. -e in *sense*, or also by phenomena from the sentential, interlexical, context, as Fr. -s in *ils, tables*. As the latter example shows, a character not having pronunciation can have (via a grapheme, see below) a relationship to another level of language than the phonemic one: Fr. -s represents the morphemic value of plural here, i.e. it has a systematic linguistic function.

Different characters are variants of a single *g r a p h e m e* if they do not differ in their systematic function (pronunciation or other, see § 0.2, fn. 11) and either each of them can occur (in any text that is not about characters) in all contextual positions in which the other can occur (stylistic variants, such as the Russian handwritten characters *đ* and *ѓ* or *Ћ* and *Ѡ* or different sets of types with typewriters and printing techniques), or each of them occurs only in such contextual positions (determined by graphemic context) in which the other cannot occur (combinatory variants, such as the two forms of Greek sigma).¹⁵ If these conditions are not met, then different characters constitute different graphemes.

A difference between two stylistic variants of a grapheme is functionally *l o a d e d* if within a single text both the variants are used without being freely interchanged (this concerns the differences between the basic characters and capitals, italics, boldface types and similar oppositions).¹⁶

A character either is *s i m p l e*, i.e. none of its proper parts is understood by the users¹⁷ as constituting itself a character, or it is *c o m p l e x*, i.e. it consists of a simple character and of one (or, exceptionally, more) accessory sign(s) of (each of) the following two kinds: (a) *d i a c r i t i c s i g n* (which is combined only with characters belonging to a specific subclass), or (b) an underlining (cf. 1.2 below), which may accompany any character.

A grapheme no variant of which contains a diacritic sign is called a *p r o t o g r a p h e m e* in the sequel, and such a constituent part of a grapheme that corresponds to diacritic signs as to its variants is called a *s u b g r a p h e m e*. Thus, e.g., the French accents (*aigu, grave, cir-*

conflexe) as well as the cedilla are subgraphemes. It may be of some interest to look for cases in which some of the variants of a grapheme are simple characters, while other are not (the French capitals in such inscriptions as LES POEMES DE P. VALERY are a mixed type, since the character E here corresponds to é, è as well as to e). A grapheme that is not a protographeme will be called a complex grapheme in the sequel.

Besides the graphemes there are other symbols each of which regularly occupies a slot in a written text; these we call the punctuation marks (comma, semicolon, colon, full stop, quotation mark and others) and (b) apostrophe, inverted commas, hyphen, etc.¹⁸

An alphabet of a language L can now be defined as a set of graphemes such that

- (a) is used to constitute written texts in L,
- (b) is complete in that it includes all the graphemes needed for any text of L, and
- (c) is minimal in that none of its proper subsets is sufficient for the texts of L.

In other words, two alphabets are present whenever a single text can be written alternatively with the use of one or another set of graphemes.

However, it should be kept in mind that the term ‘alphabet’ is often used with different meanings:¹⁹

- (i) by an alphabet one often means a set of characters rather than graphemes (e.g. whenever a table is being presented, displaying the basic types and the capitals of a given language, or its handwritten letters, and so on);
- (ii) usually, although not necessarily, an alphabet also has an (partially or linearly) ordered shape, so that the term ‘alphabet’ is also used to designate the ordering itself (e.g. in such a question as “does t precede u in the Latin alphabet?”).

Different languages often use the same alphabet. Marginal cases of the identification of alphabets can be indistinct: if the difference between the two sets of graphemes is small, it is not probable that one would speak about two different alphabets; e.g. in German the use of β is not fully obligatory, but it is more natural to speak of a single alphabet than of two (differing in that only one of them contains the quoted grapheme).

It is useful to distinguish different subclasses of graphemes within an alphabet, especially: (a) the protographemes, (b) such complex graphemes

that are ordered in the alphabet (usually immediately following the corresponding protographemes, e.g. Czech č, š, Polish Ł, German ä, ö, ü, and (c) such complex graphemes that are unordered, i.e. do not influence the order of words in dictionaries (e.g. Cz. á, ð, Fr. à).

This understanding of the term ‘alphabet’ makes it possible to state the difference between the alphabets of Latin, French, German, Czech, etc. (or between those of Serbian, Bulgarian, Russian, etc.), and also to state that Latin and English, e.g., use the same alphabet.²⁰

There are signs which differ from the diacritic signs in that they do not distinguish between characters corresponding to different graphemes and they serve only to differentiate between variants of a single grapheme; they can be combined with any character. Underlining (of various kinds) is an example of this class of signs. Often the underlining is used to mark a functionally loaded difference, and the same or a similar functional load (i.e. the same relationship either to pronunciation or to the semantico-pragmatic layers of the text) can be carried by different variants of the graphemes (italics, etc.).

Up to now we have attempted (in § 1.1) to take into account only the written text itself, without regard to its relationships to spoken language. Even so, as we noted in the introduction, and as we have seen with the definitions of the concepts “pronounce” and “grapheme”, at least the fact that some symbols have a certain relationship to the phonemes of the given language seems to be a necessary ingredient of the characterization of an alphabet. In the next section we want to complement the characterization of graphemes and their combinations with regard to further features of their relationships to phonemes and to spoken language in general.

1.2 Since pronunciation and orthography represent the core of the relationship between spoken and written texts of a language, a systematic analysis of the questions of orthography requires as one of its prerequisites an inquiry into different questions of the correspondence between graphemes and phonemes.

In a phonemic orthography the primary case of this correspondence fulfils the following conditions: (a) in any context the given grapheme is pronounced as the same phoneme, i.e. there is only one phoneme to which the grapheme corresponds,²¹ (b) in any context the given phoneme is written as the same grapheme, i.e. there is only one grapheme corresponding to the phoneme.

Condition (a) can be called the “uniqueness of pronunciation” (of a

grapheme), and condition (b) “the uniqueness of spelling” (of a phoneme).

These conditions underlie every phonemic orthography, but almost always there are also secondary cases, in which the correspondence between graphemes and phonemes is not this simple; therefore, more complex means are needed for a description of orthography and pronunciation (cf. Gak's (1962) six types of “non-absolute” graphemes).

First, often the uniqueness is violated, so that it is not possible to understand a grapheme as a function of a phoneme (with the orthography), or a phoneme as a function of a grapheme (with pronunciation). We shall speak of a relation between phonemes and graphemes, which for English includes e.g. the pairs (*/k/, k*) as well as (*/k/, c*) - i.e. the phoneme */k/* is written by means of the two graphemes — and besides the pair (*/k/, c*) also the pair (*/s/, c*) is an element of the relation, i.e. the grapheme *c* is pronounced as either of the two phonemes.

Second, it is easy to show that in the general case the correspondence concerns a pair of strings of phonemes and graphemes rather than a pair of single items: English *sh*, *ch*, etc., are regularly pronounced as a single phoneme each, and *x* corresponds to */ks/*, i.e. to a string of two phonemes; with the German *sch* or the French *eau* it is obvious that it would not be sufficient to speak of a pair of graphemes.²²

Therefore we define the relation *O r t h* (which can be interpreted as a basic ingredient of a rule of orthography as well as of a rule of pronunciation) as a pair of strings, the leftmost string consisting of phonemes, the rightmost one of graphemes; for our present purpose it seems that the length of the strings concerned will not exceed 3 (for Czech, 2), but if more complex cases are found in a language, they can be accounted for by the relation *Orth*.

Examples of the elements of *Orth* for English:
 $(/\check{s}/, sh)$, $(/\check{s}/, ti)$, $(/\check{s}/, ch)$, $(/\check{c}/, tch)$, $(/ks/, x)$.

If $(/Y/, X)$ is an element of *Orth* (the capitals being used as variables for strings of graphemes or of phonemes), then we write $X \in Sp(/Y/)$ and $/Y/ \in Pron(X)$; i.e. $Sp(/Y/)$ is the set of all strings of graphemes by means of which the phonemic string $(/Y/)$ can be spelt in the language described (according to its orthography), and $Pron(X)$ is the set of all strings of phonemes occurring as possible pronunciations of the graphemic string X .

It is convenient to restrict the definition of *Orth* in such a way that at most one of the strings constituting an element of *Orth* can be empty, so that for instance $(/\emptyset/, e)$ is an element of *Orth* in English (e.g. in *come*).

We can now define “uniqueness” more precisely than we have done above: a phonemic string $/Y/$ is spelt in a *u n i q u e* way iff Sp ($/Y/$) is a set having a single element string; on the other hand, a graphemic string x is pronounced in a *u n i q u e* way and called *u n a m b i g u o u s*, iff the set Pron (X) has a single element.

If the spelling (or the pronunciation) of a string is not unique in this sense, then the rules of the orthography (or of the pronunciation) of this string in the given language specify the conditions of choice between the different ways of spelling (or of pronunciation). If the choice of the (string of) character(s) is deliberate, not restricted by such conditions, then only different (strings of) characters functioning as variants of a single (string of) grapheme(s) are present (e.g. Russ. *ж* and *ঢ*; similarly in the case of pronunciation we find stylistic (individual, etc.) variants of a single phoneme — e.g. different sounds constituting the phoneme */r/*). Also if the choice of characters is determined by the graphemic context, one can speak about a difference in characters (variants) rather than in graphemes, cf. the definition of grapheme in 1.1.

In the literature one often speaks about an orthography being based on several principles, the main among which is the phonemic one, while the others underlie the deviations from this basic principle and can be classed more or less exactly in accordance with the levels of the language system. Often a so-called historical (or etymological) principle is distinguished, which does not belong to this strictly synchronically based patterning, but under which the sources of many deviations from phonemic writing can be subsumed (see e.g. Maslov, 1975: 307ff). Perhaps the most systematic classification of these principles (or ‘representations’) has been presented by Hall (1960). In other writings, see esp. Rahnenführer (1980: 256), the historical principle is not acknowledged at all, and such secondary cases as the different spellings of German */a:/* in *Stahl*, *Staat*, *Rabe* are characterized as instances of the morphemic principle. However, it seems important to distinguish between such cases as these (where the different spellings are conditioned by the individual morphemes, but the only motivation is determined by the historical conditions: only in a previous state was a difference in pronunciation present) and those where the different spellings correspond to different, homophonous morphemes (*mal* vs. *Mahl*, *leer* vs. *lehr-en*) or where a single string of graphemes corresponds to different strings of phonemes, but this ambiguity helps to recognize that a single morpheme is present (as is usually the case with phonemic neutralizations: German *Tod*

— *Todes*; among other cases let us recall Czech *vod-a* vs. *vod-ě*). Without distinguishing these cases it is not possible to claim that the morphemic principle as such has or provides some motivation for spellings within a phonemic orthography.

Another point on which it is useful to have more precise formulations than those using the notion of ‘principle’ concerns the spelling of G. *Rad* — *Räder*, Czech *vozu* - *vůz*, etc. It is possible to speak about morphemic factors the impact of which consists just in that the relevant orthographical rules, which are rather specific, contribute to the fact that the identity of the morpheme is rendered as overtly as possible in the given graphemic system: the protographeme of *ā* and the diacritic in *ü* point to the relationship to *a* and *o*, respectively, and thus make it as evident as possible that *Räder* and *vůz* correspond to secondary allomorphs of *Rad* and *voz-*.

A similar need for more precision can easily be found with formulations concerning other principles (syntactic, semantic, etc.). The rules based on them often are specific for small groups of instances, and it is not possible to draw a strict boundary line between a rule and an exception.²³ It is fully appropriate, as has been repeatedly stated by Vachek (1939: 103; 1942: 303), that every specific orthography is based on a single principle; the other tendencies or factors (those underlying the specific rules and the exceptions) belong to another layer, be they as important as in English, or as reduced as in Spanish.²⁴ Vachek also pointed out why so many orthographies are phonemically based: the number of phonemes in a language is so much smaller than that of morphemes (or syllables) that it is much more suitable to use the repertoire of phonemes as the set of the basic units than some more complex types of expressions (see Vachek, 1933, 293).

We can state that phonemic orthography is based on the correspondence between graphemes and phonemes, although we are fully aware of the autonomy of graphemics and of the existence of spelling rules based on another factor than the stated correspondence. The ‘principles’ can serve as cover terms for certain groups of such factors.

On the basis of the correspondence between graphemes and phonemes it is possible to specify the concept of a b i g r a p h e m e, which can be defined as a sequence of two graphemes pronounced as a single phoneme. If the bigrapheme is unambiguous (i.e., for a bigrapheme *X*, the set Pron (*X*) has a single element, see above) and exclusive (so that also Spel (Pron (*X*)) has a single element, viz. *X*), we can speak about a pure bigrapheme.

However, such a situation seems to be rather rare. Even such bigraphemes as E. *sh* or *ph* have their irregular pronunciations at morpheme boundaries (e.g. in *gasholder*, *upheaval*); see also Vachek (1933: 301f) on such an exception concerning Czech *ch* (although not in Standard Czech): Holeček's *mlachuba*. With E. *ch*, *ea* and other bigraphemes the unambiguity is clearly missing. Exclusiveness seems to be absent with E. bigraphemes.

A few notions standing close to that of a bigrapheme can now be characterized:

A *trigrapheme* is a sequence of three graphemes pronounced as a single phoneme (e.g. G. *sch*).

A *quasi-bigrapheme* can be seen in those cases where a string of two graphemes is pronounced in a way that differs from the sequence of the two phonemes corresponding to the two graphemes. More exactly, if g_i and g_j are used as variables for graphemes, $g_i g_j$ is called a quasi-bigrapheme, if an element of $\text{Pron}(g_i g_j)$ differs from the concatenation of $\text{Pron}(g_i)$ and $\text{Pron}(g_j)$, although this element is not a single phoneme.

Examples of quasi-bigraphemes:

(a) In many languages, *qu* is pronounced in the same way as *kv*; if the latter sequence is not used in the orthography of such a language, then *q* may be understood as a positional variant of *k* (or of *c*).

(b) Czech strings *di*, *ti*, *ni*, pronounced as /d+i/, etc, correspond to a specific spelling rule; for *dě*, *tě*, *ně* such a rule can be formulated on the basis of the "neutral" *de*, *te*, *ne*; however, *dy*, *ty*, *ny* do not represent a fully neutral basis for the former triple, since a specific spelling rule states the necessity to spell /i/ as *y* after *d*, *t*, *n*.

Another case (in a certain sense symmetric to a bigrapheme) is that of a grapheme pronounced as a string of two phonemes: *x*, Greek theta, phi, khi, psi, ksi. Also Czech ě corresponds to two phonemes after *m*, *p*, *b*, *v*, *f*.

2.1 Since, as we have seen, the correspondence between phonemes and graphemes is crucial for a phonemic writing system (even if autonomous status of the written norm is fully taken into account) and the orthography is accessible to conscious interventions in a degree to which rules of grammar (and phonemics) are not (see Sect. 0.2), it is quite understandable that often requirements concerning the appropriateness of an orthographic system are discussed. These requirements concern first of all the relationships between phonemes and graphemes, i.e. our relation Orth, and their

relative fulfilment is relevant for the practical appropriateness of a given orthographic system.

The main requirements were formulated by Vachek (1964: 125) and can be summarized (with modifications) as those of

- (i) transparency: the written form should be easily convertible into a string of morphemes, the path from the graphemic form of the text to its meaning should be straightforward;
- (ii) learnability: the rules of spelling and pronunciation should be simple and general.

These requirements also underlie the degree of the uneasiness of orthographic rules, their acquisition and the corresponding practice.

Due to the secondary character of graphemics, there is a certain asymmetry: in the process of language acquisition the usual case is that when the rules of spelling and of pronunciation are to be learned, the spoken language (in its basic layer) is already known. With spelling, the exclusive graphemes certainly correspond best to the requirements. However, with pronunciation the unambiguity of graphemes is not such a compelling criterion. For the reader, the written text is the source of information concerning all the other levels of the structure of language, so that it is not always crucial whether the pronunciation is determined in the easiest way or not. A simple identification of the morphemes is equally important for the reader, and so is an easy access to the units of still deeper levels of the system of language. Therefore, if a certain occurrence of a grapheme corresponds to different phonemes only in different graphemic or morphemic contexts, this should not be understood as causing difficulties.

However, due to the basic one-to-one prototype of the relation Orth (i.e. of the relationship between phonemes and graphemes), such a case of ambiguity is necessarily accompanied by a case of non-exclusivity. Since no phoneme seems to be limited to a single morpheme (nor to a narrowly restricted subclass of morphemes), whenever a grapheme g_i is ambiguous, so that $\text{Pron}(g_i)$ has two elements, say / g_i / and / f_i / (if these symbols are understood as variables for phonemes), then g_i cannot be exclusive without idiosyncratic rules being present: if g_i is exclusive, then there are idiosyncratic rules for its pronunciation as / f_i / or / f_j / (if these are not mere morphophonological variants, then their distribution hardly can be determined by graphemic context or by oppositions between morphemes). Thus, ambiguity can be combined with exclusivity only when its solution is rather complex (such may be the case in the absence of graphemic means for rendering voicedness with some occlusives in e.g. Hittite or Mycenaean, or,

the E. bigrapheme *th*).

Therefore, in a case of ambiguity determined by general rules, such as that concerning e.g. Czech /t/, /d/, /ň/, exclusivity cannot be expected. The graphemes *t*, *d*, *n* are pronounced as /t/, /d/, /ň/ only before *i* and *ě*, i.e. their ambiguity in the given point is solved by the graphemic context. As our preceding formulation predicts, these graphemes are not exclusive: in other graphemic contexts the given phonemes are spelled by *ť*, *đ*, *ň*: *ťal*, *lod*, *oceňme*.

The lack of exclusiveness can be said to be harmless for the process of reading, but it causes difficulties in writing, since here it is necessary to make decisions concerning the choice of spelling, the contextual criteria involving in some cases only the adjacent phonemes or graphemes — e.g. the E. string /šn/ is spelled *tion*, rather than *shn*, *chen*, etc. — and in other cases lists of morphemes (which have to be consulted e.g. for the choice of spelling of, for example, *mite* vs. *mite* or *mice*). It is possible to compare the conditions and requirements of writing with those of reading and recall, for instance, that reading is more frequent (so that its requirements, including the usefulness of avoiding homographs, cannot be overlooked), whereas writing is more troublesome in that its output has to be learned (even by native speakers, who have acquired the spoken norm without being specifically taught), and thus every issue complicating writing weighs more than a comparable complication of reading (since the possibility of reading errors is diminished by the fact that it is much easier to check the output of reading than that of writing). An appropriate orthographic system certainly should allow for writing to have the character of a mechanical process, determined by general rules and containing no necessity of choices based on criteria concerning sentential context. Especially, writing down a word form in a spoken utterance should not require waiting for another word-form that can be indefinitely distant; such a difficulty e.g. may occur with a case of subject verb agreement, if the subject can be placed at the end of the sentence, cf. Czech *Přišly nás tehdy, sotva jsme se nastěhovali do města, kde ..., navštívít Marie a Karla* — “(There) came (Plur. Fem.) to visit us then, just after we moved to the town where..., Marie and Charlotte”, where the ending /i/ in the first word would be spelled *i* if *Karel* (*Charles*) rather than *Karla* (*Charlotte*) were the last word of the utterance, since the difference in gender is rendered by the written form of the ending and not by its phonemic shape.

In any case, an appropriate orthographic system should contain as few

hindrances as possible, whether they are more immediately connected with writing, or with reading. All such complications make the acquisition of an orthographic system and of the skills for its use more difficult than necessary. However, it is true that the complications can be classified along the degree of inappropriateness they carry. Some of them bring difficulties in one respect, but are useful in another (esp. if they facilitate the identification of morphemes in reading), others are more or less tedious. Since the complications, i.e. the deviations from the one-to-one correspondence between phonemes and graphemes concern two dimensions, viz. that of complexity and that of univocality, we present in the following section two tentative scales for such a classification. The scales can be made more precise in the future if more factors are taken into account (e.g. also the complexity of the string of phonemes in the elements of the relation Orth, or the degrees of difficulty of the criteria for disambiguation should be reflected).

2.2 The first of the scales we would like to present (with many simplifications) concerns the complexity of the graphemic string of an element of the relation Orth. Before formulating the scale we want to define one more concept that will be needed:

A subgrapheme is called *r e g u l a r* if it corresponds, in the relation Orth, to a phonemic correlation; thus in Czech e.g. the diacritic sign is regular with *á*, *í*, *é*, etc. since the corresponding phonemes are marked as the long counterparts of /a/, /i/, /e/; similarly, also in *š*, *ž*, *č* the subgrapheme is regular, corresponding to a correlation between the phonemes /š/ and /s/, etc., the opposition of which is that of compactness vs. non-compactness. The opposition between /ř/, /đ/, /ň/ and /t/, /d/, /n/ probably has the same character, so that in *ř*, *đ*, *ň* we again have a regular diacritic sign. Perhaps this is not the case with *ř*, although the morphophonemic alternation /r/ - /ř/ is quite parallel to that of /t/ - /ř/, etc., and the spelling rules for *ř* are fully phonologically based (the difference between the voiced and unvoiced combinatorial variants of /ř/ is not reflected in the graphemics).

(A) The scale of complexity (of the graphemic string of an element of Orth):

- (i) a protographeme
- (ii) a complex grapheme the subgrapheme of which is regular
- (iii) a complex grapheme with an irregular subgrapheme
- (iv) a string of protographemes (having more than one element)

- (v) a string some of whose elements are complex, but with no irregular subgraphemes included
- (vi) a string some of whose elements have an irregular subgrapheme
- (vii) as in (vi), with the irregular subgrapheme corresponding to one more phoneme in the pronunciation (or with another difference between the number of graphemes and that of phonemes)

(B) The scale of univocality is formulated here first of all according to the degree of univocality of the operation Sp, since — as we have seen in Sect. 2.1 — the rules of spelling usually are more difficult to learn than those of pronunciation (the latter being more easily inferred by a speaker who is already acquainted with the spoken norm):

- (a) absolute bi-uniqueness
- (b) relative bi-uniqueness (the specific rules concerning deviations being put aside...)
- (c) regular deviations: different graphemes are used for a single phoneme with different morphemes, the choice being given by a general rule and being corroborated by phonemic alternations in the given morphemic position (with a different phonemic context)
- (d) as in (c) without the corroborating alternations
- (e) irregular deviations, the choice being given by a difference between lexical morphemes
- (f) irregular deviations with the choice given by a morpheme whose single graphemic shape is thus ensured
- (g) irregular deviations with the choice given idiosyncratically (traditionally, without a functional justification)

As Table 1 illustrates, using examples from Czech, the scale (A) seems mostly to be of little relevance for the appropriateness of an orthography; it is just the degree (vii) that appears to be important there; on the other hand, the scale of univocality displays different degrees corresponding to the degrees of easiness of individual points of the orthography and of its appropriateness for writing as well as for reading. More detailed comments on the issues of Czech orthography illustrated in the table are presented in Sgall (1986) (written in Czech, since it is mainly of interest to specialists in Slavistics).

The graphemes *y* and *ě* are not characterized as such in the table, but only in the strings or in examples of specific positions in which they occur; examples are also given for the irregular occurrences of *z*, *s*, *i*, *a*. Similar

		bi-uniq.	rel. bi-un.	regul. alt.	differ.	unifc.	idiosyncr.
Scale B	A	(a)	(b)	(c)	(e)	(f)	(g)
(i) simple		<i>b, f, g,</i> <i>k, l, m, p,</i> <i>r, v</i>	<i>a, c, d, e,</i> <i>h, j, n, o,</i> <i>s, t, u, z</i>	<i>psy</i> <i>byl</i>	<i>-yně</i>	<i>větší, jazyk,</i> <i>zpěv, kurs,</i> <i>stadium,</i> <i>aktivum</i>	
(ii) diacr. reg.		<i>é,</i> <i>š, ž</i>	<i>á, í, ó, ú,</i> <i>đ, ñ, ī, č</i>		<i>být</i>		
(iii) d. irreg.		<i>ř</i>	<i>ü</i>				
(iv) string			<i>di, ni, ti,</i> <i>dy, ny, ty</i>		<i>objetí</i>		
(vi) str. diacr.		<i>dě, ně, tě</i>			<i>temnější</i>		
(vii) excessive phoneme		<i>pě, árie</i>	<i>x, město</i>				

Table 1.

examples can also be added for other graphemes (e.g. *ředkev*). In the examples the grapheme concerned is underlined.

Cases (v) and (d) seem to be absent from Czech orthography.

The table certainly does not cover every issue important for a characterization of an orthographic system. We do not analyze here questions of punctuation, of writing capitals and compound words, of hyphenating. Also the complicated questions connected with the spelling and pronunciation of loan words are not discussed in the present paper.

The table can help in characterizing or classifying the orthographic rules, as far as they specify the operations Sp and Pron for the given orthographic system. It can be seen that in Czech the rules necessary for a specific

cation of the spelling of such phonemes as /b/, /f/, /g/, /k/ are general, whereas for /a/, /c/, /d/, /e/, etc. there also exist specific rules or small classes of exceptions. Some of these serve the differentiation of homophonic morphemes, be it regular grammatical morphemes — (c), (d), or lexical units (e); other specific rules cause similar difficulties, complicating the acquisition and use of the orthography, but their functional impact either is limited to the graphemic unification of a morpheme (f), or is fully absent from the synchronically viewed system of the given language, the rules being idiosyncratic and determined just by tradition and/or by an imperfection of the codification.

3. We may conclude that the development of the approaches to graphemics has shown that the graphemic level has its autonomous existence, but even so the relationship between this level and that of phonemics has to be reflected in a systematic description of graphemics. This relationship is asymmetric, graphemics being secondary, less close to the core of the linguistic system. The relationships between graphemics and morphemics, syntax, the level of meaning and the structuring of text have an importance of their own (especially in those domains which were not in the centre of attention in this study: the use of capitals, rules concerning word boundaries, punctuation, etc.), although they are in principle more peripheral than the correspondence between graphemes and phonemes.

This viewpoint can well be used, as we have seen, for the aims of an investigation of the appropriateness of an orthography. Even those who are skeptical about a purely theoretical evaluation of orthography admit that for practical purposes it might be an advantage to bring the graphic system into harmony with the phonic system. This is the standpoint of Uldall (1944), who also states that "the orthography is often more conservative than the pronunciation, and has further been artificially tampered with by printers and grammarians." However, the practical character of orthographic conventions makes it natural to also require a theoretically oriented investigation of the questions of the appropriateness of orthographies, paying due respect to the conditions of writing as well as to those of reading. As Vachek and others have convincingly shown, one of the relatively most useful aspects of deviations from the grapheme-to-phoneme correspondence is that of the identity of morphemes, especially in those cases where two homophonic morphemes are differentiated graphemically.

Two remarks seem to be in place in this connection:

First, such deviations from the phonemic basis of the orthography are usually connected with so-called historical (etymological) spelling. It is relatively rare that such a graphemic differentiation of morphemes comes into existence, be it by a deliberate decision (cf. Ital. *è* - "is" vs. *e* - "and", *tè* - "tea" vs. *te* - "you-Acc., Dat. Sing."), or without it (cf. Bolinger, 1946, on such a bifurcation as *passed — past*, *complement — compliment*); much more often two morphemes that have merged in their phonemic shape maintain different etymological spelling (cf. E. *write* and *right*, *be* and *bee*, *week* and *weak*,...).

Second, it seems not to be the usual case that in maintaining a historical spelling one of the main criteria consists in a functional evaluation of whether the etymological difference facilitates the identification of a morpheme (or even the differentiation of more morphemes). The development of orthographies has been biased by most different sociolinguistic factors and does not directly substantiate the view that the orthographic stability of morphemes plays a decisive role in the system of language. Most often the historical (etymological, or just traditional) orthographic deviances are maintained in both cases, be they morphemically loaded, or not. Also, Berry (1958), e.g., who investigates the requirements on newly coined alphabets and quotes many publications with different standpoints on the relevant questions, writes (p. 754) that "it is also agreed that the orthography should be such that one who knows the pronunciation of a word should be able to spell it correctly", thus denying the usefulness of graphemic unification and differentiation of morphemes.

We have presented an analysis of the theoretical questions of graphemics, which might be useful as a basis for discussions of the usefulness of an orthographic reform, since it is advisable not to base proposals for such a reform on just subjective impressions arising from onesidedness. It has not been among our objectives to discuss such proposals, the acceptability of which is always conditioned not only by linguistic inquiries, but also by sociological conditions. It seems that for such a discussion a linguistic inquiry can be useful in three directions:

- (i) Are there deviations from the grapheme-to-phoneme correspondence that substantially diminish the degree of the appropriateness of the given orthography?
- (ii) Is it possible to state the proposed change in the orthography in simple, general terms (i.e. would the new rules be easy for the adult speakers who have been acquainted with the current system)?

(iii) Would the reform make access to the "old" literature relatively difficult for the coming generation (trained already in the new system)?

We wanted to contribute to points (i) and (ii); point (iii) has not yet found fully appropriate methods to be studied systematically, but it can be evaluated at least in part according to the results of (ii). We are aware that there are also questions which cannot be solved by linguistic means; this concerns especially the social conditioning of a reform: e.g. for a language the native speakers of which constitute dozens of independent countries such a task is much less realistic than for a language mostly concentrated in a single country.

Since a spelling reform itself is not among the objectives of the present study, we leave these questions open. What we wanted to achieve is to deepen the theoretical insight into the level of graphemics and its position in the system of language, to handle consistently the necessary concepts applied in the characterization of graphemics by many scholars from Baudouin to Váček, Penttilä, Haas, Bierwisch, Ruszkiewicz and others. We are convinced that such an insight is indispensable not only for practical purposes, be they connected with spelling (and pronunciation), or with an automatic treatment of written texts, but also for a complete theoretical description of a language, since the written norm is an integral part of modern language, equally essential to it as its spoken norm.

NOTES

1. This is similar to the situation in classical phonemics, where (see Trubetzkoy, 1939, p. 32, on Germ. /h/ and /ŋ/), to a certain degree, the closeness of phonetic properties has to be taken into account in the first phase of analysis. In graphemics, criteria concerning phonemics are necessary, even though the main method for distinguishing between different graphemes is that of distributive segmentation and of checking minimal pairs, as Piirainen (1971) duly notes; the exclusive use of elementary set-theoretical means known esp. from Penttilä (1971 and the writings quoted there) without a functional criterion would not be sufficient for an identification of capitals and other variants as belonging to a single grapheme. This also concerns one of the most systematically elaborated approaches to graphemics as an autonomous system, derivable — on a par with phonemics — from a neutral underlying structure, which was presented by Allén (1967 and his previous publications quoted there; see the comments by Hellberg, 1974: 35ff). Allén works with what can be called distinctive features of graphemes (big loop in *o*, small lower loop in *a*, small upper loop in *e*, diacritic marks, etc.); however, it remains open to discussion whether these features do not belong (also) to the level of characters; cf. also Harweg (1971) and Brekle (1971).

2. Váček (in press) shows to what extent also the rules for using capitals in the orthographies of different languages reflect (indirectly) structural properties of these languages as func-

tioning systems.

3. As Derwing and Dow (1984) remark, knowledge of orthography may be a deciding factor (in the form of ‘careful pronunciations’) in phonetically ambiguous cases, where there are no morphemic cues. This confirms the view that an autonomous approach to graphemics is well compatible with regarding the correspondence between graphemes and phonemes as the prototypical relationship.

4. Vachek (1942: 266ff), as well as Gak (1962), McIntosh (see Heller, 1980: 87) and others speak about the phoneme as constituting the meaning of the grapheme, which makes it possible to use such terms as synonymy and ambiguity with respect to the relations between graphemes and phonemes. However, as Vachek (1972) remarks, quoting Haas, the term ‘phonemic meaning’ should better be replaced by that of correspondence. (Let us note that also in such relationships as between a sound and a phoneme, or a morph and a morpheme, a morphemic case and a syntactic relation, etc., the term ‘meaning’ does not seem to be appropriate, although it is possible to speak about means and function here.)

5. In the prototypical case the grapheme-to-phoneme correspondence is determined by general rules, so that a “detour” would consist in ignoring these rules rather than in describing the written norm as dependent on phonemics. In English and French the irregularities are much more frequent than in Spanish, German, Czech, Russian and many other languages; some of them are known to be substantiated by morphemics (see esp. Bolinger, 1946) in maintaining the identity of morphemes (*-ed*) or their difference (*write — right*), but in many cases either no such synchronic substantiation is present (*keel — veal*), or even the graphemic shape disguises a morphemic and phonemic difference (*read Pres. — read Pret.*).

6. Therefore also the users’ attitudes play a more important role with respect to the identification of the units of the graphemic level than they do in other domains of language; this concerns e.g. such questions as what is understood as a diacritival mark, see Sect. 1.1 below.

7. The rules specifying the phonemic representation of these morphemes are not relevant for the orthography, but they are important for the pronunciation, e.g. to identify the difference in the pronunciation of the endings in *landed* and *crossed*.

8. Sedláček (1977: 48f) points out that not only (surface) syntactic criteria, but also those from the level of meaning are present here (the latter include, as we are convinced, also those from the domain of the topic-focus articulation, understood by Sedláček as a specific level).

9. However, it is not exact to say (as e.g. Edgerton, 1941 does) that e.g. English writing (in such cases as *ant* vs. *aunt*) “suggests the ideas directly,” rather than the morphemes; even most abbreviations, including \$ or 2nd, correspond to language specific morphemes (although they can spread from one language to others); a direct connection with a structuring of the cognitive content that is not language specific can be seen in the symbolic systems of mathematics ($3^2 = 9$; $x(Rx)$; αE , chemistry (H_2O), etc.

10. Among the two main turning points in the development of writing, the first one — the beginning of phoneticization — occurred when this parallelism emerged (as a peripheral phenomenon; the main principle of the writing system still was that of a graphical representation of morphemes); this happened with different cultures in different epochs. The second stage is that of phonemic writing, achieved by the Greeks, with the parallelism as the prototypical principle. In between there are the systems based on syllabic and consonantal principles (with a partial parallelism), and various systems of mixed kinds.

11. An open question concerns the difference between (i) such functions of graphemes as

those in Germ. *viel* and *fiel* or in French *halète* and *altère* and (ii) such uses of variants of graphemes as in Germ. *gehen* vs. *Gehen*, E. *robin* vs. *Robin*. In case (i) a functional difference is present that in the unmarked, prototypical case (where ambiguity is absent) is rendered by different phonemic means; in the sequel these functions will be called "systematic" and will be used in the definition of 'grapheme'. In case (ii) it is typical that such differences as those between proper names and other nouns, between nouns and other parts of speech (as well as between the initial words of a sentence and other words) usually are not rendered by phonemes, so that the orthographic conventions here are not so closely connected with the core of the language system; we understand these cases as different kinds of functional loading of variants of phonemes, although we are aware that this loading differs substantially from the usual stylistic values of variants in phonemics or morphemics.

12. For a generative approach to Czech phonemics the idea of a hidden phoneme, partly reflected by /j/, would be important, see Poldauf and Šprung (1968), cf. also Hronek (1971) and the literature quoted there; Barnetová's (1975) analysis is valuable for the generative phonemics of Russian and for a comparison of generative phonemics with Praguian morphophonemics. A detailed and systematic treatment of graphemics on the background of generative phonemics was presented by Hellberg (1974) for Swedish. We use '/ ' to denote the classical phoneme. It should be determined by appropriate psycholinguistic methods whether (as we assume) the classical phoneme also plays a major role in the acquisition of writing systems by children. It is not exact to say that the natural support for learning is the sound, whereas the phoneme would be an abstract construct; rather, the classical phoneme seems to be natural (directly accessible to a native speaker's intuition), the differences between combinatory variants (allophones) not being always consciously identified; the abstract phonemes or morphophonèmes seem to have the character of constructs (useful within a theory). Thus e.g. a Czech child does not necessarily become aware of the existence of such sounds as the velar *γ* (a combinatory variant of /n/ or the unvoiced *R̄* (a combinatory variant of /t/), but s/he necessarily learns that e.g. the phoneme /t/ should be written *d* in such a position as in *pojte*, to say nothing of such graphemic and phonemic differences as those among *k*, *c*, *č* in *matka*, *matce*, *matčin*.

13. These notions are understood here in the sense of classical Praguian phonemics, as was stated in Sect. 0.3.

14. This formulation should be made more precise to show that e.g. French *h* is a grapheme, whereas the hyphen is not: The hyphen is used within the graphemic shapes of compounds of certain kinds and as an indication of the division of a word-form in two lines, but in neither of these used does a hyphen render a fixed part of a simple word-form.

15. Gak (1962) distinguishes between positional, functional, individual and notional variants of graphemes, pointing out that the shape of characters itself is not always decisive, such a classification being functionally based.

16. Different kinds of functional loading should be distinguished: the various uses of italics or boldface types, further marking the beginning of sentences and marking proper nouns (and, as in German, nouns) constitute at least three layers, each of which differs from what we have called 'systematic functions'.

17. It is not trivial to explain why this is so with the Russian *u* (as compared with *u*, although not with French *ç*. Polish *ę*, *ą* can be understood as complex, and the same probably holds of *t*, the diacritic mark of which does not occur with any other grapheme. In Russian *u* has the same element as *u* (from a purely graphic viewpoint), but its relationship to *u* — a vowel — is of quite a different character than that of *u* to *u*, due to the substantially greater difference in pro-

nunciation. This shows that even with such extrinsic layers as the alphabets the functional view of the relationship to other layers of language is essential. A discussion of the evaluation of diacritics was presented by Berry (1958: 761).

18. Here again the functions rather than the shapes of the symbols are relevant for their classification; in principle there is no reason why e.g. the apostrophe could not be used as a character in an alphabet (this is the case e.g. in some of the transliteration systems substituting the Latin alphabet for the Russian one). As Sedláček (1977: 46) points out, the rules concerning the comma are first of all syntactico-semantically (rather than phonetically) based not only in German (see Baudusch, 1980), but also in Czech, and even in English (cf. Quirk *et al.*, 1974: 1055). Other non-graphemic, although orthographical rules concern the hyphen, the blank, etc. (as for German, see now esp. Hofrichter, 1980, on the morphemic and other correspondences reflected by the rules of hyphenation).

19. The term alphabet is also used in connection with syllabic and similar writing systems (cf. e.g. Istrin, 1965). Moreover if e.g. *q* is properly understood as a variant of *k* in various European languages, than an alphabet in the usual meaning of the term contains not only graphemes, but also their positional variants.

20. It would be possible to go further and include also e.g. the differences concerning the use of capitals and of punctuation marks in the definition of an alphabet, but this does not seem necessary in the context of this study.

21. In the case of a phonemic neutralization (e.g. of /d/ and /t/ in Germ. *Tod — Todes*, Cz. *hrad — hradu*) we speak — with a certain simplification — of an identity of the phoneme (/d/) within a single morpheme, leaving here aside subtle problems of phonemic theory, since we consider it unmarked within a phonemic orthography to use a single grapheme for the neutralized as well as for the basic shape of the phoneme (we consider e.g. Lat. *scripsi* to constitute a phonetically oriented deviation from a phonemic orthography).

22. We do not consider such a string as E. *sh* (or G. *sch*) a single grapheme, since it seems to be unnecessarily redundant to speak about three graphemes: *s*, *h* and *sh*; other arguments for such a standpoint were presented by Magnusson (1972, p. 25). The string *sh* is distinguished just by being an element of the domain of the operation Pron, see below. (Even in such a case as Cz. *ch*, which is listed as a member of the alphabet, we hesitate to join Bazell, 1956, and to speak about a grapheme, which would then consist of two other graphemes.) The elements of the relation Orth correspond closely to Heller's (1980, 95ff) phonographemes; cf. also Hořejší's (1969) concept of graphoneme. It would also be possible to say that *s* before *h* is pronounced as /ʃ/, i.e. that *h* is relevant only as a contextual item; however, in some cases — e.g. with Fr. *au*, *eau* — it is not possible to tell which grapheme should be classed as the pronounced one, and it would not be easy to draw the boundary between the two classes of graphemic strings. Therefore we prefer to understand the entire strings as being pronounced in all cases, although we realize that this viewpoint can be made more precise.

23. Cf. Derwing, Priestly and Rochet (this volume) on E. *ea*, pronounced either as /i/, or as /e/.

24. If this fact is not respected, i.e. the 'principles' are understood as belonging to the same layer, and if, moreover, the historical factors are not understood as specific, then it is too easy to overlook the fact that e.g. the spelling of Germ. *Kaiser*, *Achse*, etc. (which can be delimited in terms of lexical morphemes) has no systematic basis in the synchronic system of the language, representing just a surviving, although isolated residue, which unnecessarily complicates the orthographic system.

BIBLIOGRAPHY

- Allén, S. 1967. "Lingual expression and generative grammar." In: *Actes du X^e congrès internat. des linguistes* (Ed. by A. Graur *et al.*), Bucarest 1967, 235-239.
- Barnetová, V. 1975. *Morfologie slovesného vidu v ruštině* [The morphemics of verbal aspect in Russian]. Prague 1975.
- Baudusch, R. 1980. "Zu den sprachwissenschaftlichen Grundlagen der Zeichensetzung." In: Nerijs and Scharnhorst (1980), 193-230.
- Bazell, C.E. 1956. "The grapheme," *Litera* 3, 43-46; reprinted in and quoted from Ruszkiewicz (1981), 66-70; also in Hamp *et al.* (1966), 359-361.
- Berry, J. 1958. "The making of alphabets." In: *Actes du VIII^e congrès internat. des linguistes* (Ed. by E. Sivertsen), Oslo, 752-764.
- Bierwisch, M. 1971. "Schriftstruktur und Phonologie." ASG-Bericht 11, ZISW DAW, Berlin; printed in *Probleme und Ergebnisse der Psychologie* 43, 1972, 21-44.
- Bolinger, D.E. 1946. "Visual morphemes," *Language* 22, 333-340; reprinted in Ruszkiewicz (1981), 45-56.
- Brekle, H.E. 1971. "Einige Bemerkungen zur Graphematik-Diskussion," *Linguistische Berichte* 16, 53-59.
- Chomsky, N. and M. Halle. 1968. *The sound pattern of English*. New York.
- Derwing, B.L. and M.L. Dow. 1984. "Orthography as a variable in psycholinguistic experiments," this volume.
- Derwing, B.L., Priestly, T.M.S. and B.L. Rochet. 1984. "The description of sound-spelling relationships in English, French and Russian: Progress, problems and Prospects," this volume.
- Deschamps, A. 1982. "L'orthographe de l'anglais: Cohérence ou fatras." In: *Grammaire transformationnelle*, edited by J. Guéron and T. Sowley, Paris, 269-289.
- Edgerton, W.F. 1941. "Ideograms in English writing," *Language* 17, 148-150; reprinted in Ruszkiewicz (1981), 40-44.
- Gak, V.G. 1962. "Orfografija v svete strukturnogo analiza," In: *Problemy strukturnoj lingvistiki* (Ed. S.K. Šaumjan), Moscow, 206-221.
- Hall, R.A., Jr. 1960. "A theory of graphemics," *Acta linguistica* 8, 13-20; reprinted in Ruszkiewicz (1981), 71-80.
- Hamp, E.P., Householder, F.W. and R. Austerlitz (Eds.) 1966. *Readings in linguistics II*, Chicago and London.

- Harweg, R. 1971. "Buchstabe und Graphem," *Ling. Berichte* 13, 78-80.
- Hellberg, S. 1974. "Graphonomic rules in phonology," *Nordistica Gothoburgensis* 7, Gothenburg.
- Heller, K. 1980. "Zum Graphembegriff." In: Nerius and Scharnhorst (1980), 74-108.
- Hofrichter, W. 1980. "Die geltende Regelung der graphischen Worttrennung (Silbentrennung) und Ansätze zu ihrer Vereinfachung," In: Nerius und Scharnhorst (1980), 109-139.
- Hořejší, V. 1969. "Formes parlées, formes écrites et systèmes orthographiques des langages," *Folia linguistica* 5, 185-193.
- Hronek, J. 1971. "K otázkám hláskosloví obecné češtiny" [On the questions of the sound pattern of Common Czech], *Slavica Pragensia* XIII (AUC - Philologica 2-3), 227-239.
- Istrin, V.A. 1965. *Vozniknovenie i razvitiye pis'ma*. Moscow.
- Luedsorff, Philip A. 1986. *Constraints on error variables in grammar: Bilingual misspelling orthographies*. Amsterdam: John Benjamins.
- Magnusson, W.L. 1972. "The orthographic code of Spanish," *Linguistics* 82, 23-51.
- Marcel, T. 1980. "Phonological awareness and phonological representation: Investigation of a specific spelling problem." In: *Cognitive processes in spelling* (Ed. by U. Frith), New York, 373-403.
- Maslov, Ju.S. 1975. *Vvedenie v jazykoznanie*. Moscow.
- Nauclér, K. 1983. "Connections between spoken and written language: Evidence from three investigations on normal and pathological written performance," *Journal of pragmatics* 7, 595-602.
- Nerius, D. and J. Scharnhorst (Eds.) 1980 *Theoretische Probleme der deutschen Orthographie*. Berlin.
- Penttilä, A. 1971. "Zur Grundlagenforschung der geschriebenen Sprache," *Ling. Berichte* 16, 49-52.
- Piirainen, I.T. 1971. "Grapheme als quantitative Größen," *Ling. Berichte* 13, 81-82.
- Poldauf, I. and K. Šprung. 1968. *Čeština jazyk cizí* [Czech as a foreign language], Prague.
- Pulgram, E. 1951. "Phoneme and grapheme: A parallel," *Word* 7, 15-20; reprinted in Ruszkiewicz (1981), 57-65.
- Quirk, R., Greenbaum S., Leech, G. and J. Svartvik. 1974. *A grammar of contemporary English*. London.
- Rahnenführer, I. 1980. "Zu den Prinzipien der Schreibung des Deutschen."

- In: Nerius and Scharnhorst (1980), 231-259.
- Ruszkiewicz, P. 1976. *Modern approaches to graphophonemic investigations in English*. Katowice.
- Ruszkiewicz, P. 1981. *Graphophonemics. A book of readings*. Katowice.
- Sedláček P. 1977. "Úvahy o reformě německého pravopisu" [Inquiries into a reform of German orthography], *Naše řeč* 60, 47-48.
- Sgall, P. 1986. "Charakteristika českého pravopisu," *Explizite Beschreibung der Sprache und automatische Textbearbeitung* XII, 47-100.
- Steinberg, D.D. 1972. "On Chomsky and Halle's optimal orthography," *Working papers in linguistics* (Honolulu), 85-164.
- Stetson, H. 1937. "The phoneme and the grapheme." In: *Mélanges de linguistique et de philologie offerts à Jacq. van Ginneken*, Paris, 353-356; reprinted in Ruszkiewicz (1981), 35-39.
- Stuart, D.G. 1958. Contribution to the discussion of Berry's paper, *Actes du VIII^e congrès internat. des. linguistes* (Ed. by E. Sivertsen), Oslo, 767-768.
- Tauli, V. 1977. "Speech and spelling." In: *Advances in the creation and revision of writing systems*, edited by Joshua A. Fishman, The Hague, 17-35.
- Trubetzkoy, N.S. 1939. *Grundzüge der Phonologie*, TCLP 7, Prague.
- Uldall, J.J. 1944. "Speech and writing," *Acta linguistica* 4, 11-16; reprinted in Hamp et al. (1966), 147-151.
- Vachek, J. 1933. "Český pravopis a struktura češtiny" [Czech orthography and the structure of Czech], *Listy filologické* 60, 287-319.
- Vachek, J. 1942. "Psaný jazyk a pravopis" [Written language and orthography]. In: *Čtení o jazyce a poezii* (Ed. by B. Havránek and J. Mukařovský), Prague, 229-306.
- Vachek, J. 1942. "Psaný jazyk a pravopis" [Written language and orthography]. In: *Čtení o jazyce a poezii* (Ed. by B. Havránek and J. Mukařovský), Prague, 229-306.
- Vachek, J. 1948. "Written language and printed language," *Recueil linguistique de Bratislava* 1, 67-74; reprinted in Vachek (1976), 121-126.
- Vachek, J. 1949. "Some remarks on writing and phonetic transcription," *Acta linguistica* 5, 86-93; reprinted in Hamp et al. (1966), 152-157; also in Vachek (1976), 127-133.
- Vachek, J. 1959. "Two chapters on written English," In: *Brno Studies in English* 1, 7-34; reprinted in Vachek (1976), 408-441.
- Vachek, J. 1964. "On some basic principles of 'classical phonology,'"

- Zeitschrift f. Phonetik, Sprachwissenschaft und Kommunikationsforschung* 17, 409-431.
- Vachek, J. 1972. "The present state of research in written language," *Folia linguistica* 6, 47-61; reprinted in Vachek (1976), 134-146.
- Vachek, J. 1976. *Selected writings in English and general linguistics*, Prague.
- Vachek, J. (in press): "Funkce velkých písmen v dnešní angličtině a v našich jazyčích [The function of capitals in Modern English and in our languages]," to be published in *Prague studies in English*.
- Volockaja, Z.M., Mološnaja, T.N. and T.M. Nikolajeva. 1964. *Opyt opisanija russkogo jazyka v jego pis'mennoj forme* [A study of the description of the Russian language in its written form]. Moscow.
- Weigl, E. and M. Bierwisch. 1970. "Neuropsychology and linguistics: Topics of common research," *Foundations of language* 6, 1-18.

APPENDIX

We have already quoted Berry's important requirement: "the orthography should be such that one who knows the pronunciation of a word should be able to spell it correctly." A similar requirement was formulated also by others, now see esp. Tauli (1977, quoting A. Hill), who presents convincing arguments for spoken language to be understood as primary in its relationship to written language; he subscribes to Hultman's statement that theoretical linguistic speculations should not lead to widening the split between the two systems in absurdum and that if writing does not give clues to pronunciation, we get as the result wrong pronunciation and social complication. Tauli also notes that phonemic orthography is the most economic system, which is the easiest in learning to read and write, since it employs the fewest symbols and rules to represent speech. Thus every deviation from the phonemic principle must be carefully calculated; when Chomsky and others see an advantage in having e.g. the same vowel grapheme in *write* and *written*, they take notice "only of some morphophonemic spellings in English and some other languages, ignoring the opposite spellings," such as *profound:profundity*, *hale:health*, *lead:led*, *crisis:crit-ical*, *delude:delusion* (p. 25); such a reading problem as that caused by the English orthography, which puts an intolerable and traumatic burden on learners, is unknown in Finnish, Estonian and other languages with phonemic orthography (p. 28f).

The fact that the present English orthography does not make it possible to predict the spelling of a given phonemic string, is also stated by Deschamps (1982, 288), who presents a detailed account of "the logical correspondence" between the graphemic level and a level of generative phonemics.

Chomsky and Halle's claim that an orthography based on their underlying phonological representations would be optimal was critically analyzed already by Steinberg (1972), who found (on the basis of an experiment with 25 native speakers) that the 'vowel shift rule' is not productive in Modern English, and pointed out that it would be uneconomical to store lexical units in the supposed abstract forms. Instead, Steinberg proposes a surface phonemic representation as psychologically real (and definable in the generative framework); this representation, which has much (see p. 96) in common with structuralist phonemics, is claimed to constitute a suitable basis for an orthography that would be easy to learn.

We may conclude that the papers quoted in this Appendix support the treatment of orthography as a secondary component of grammar (as for arguments from neurolinguistics, see Luelsdorff, 1986, § I.3), and corroborate two other basic assumptions accepted above:

- (i) Structural (Pragian) phonemics offers a more realistic basis for comparative studies of the phonemic and graphemic levels than Chomskyan underlying phonological representations.
- (ii) The autonomous written norm is secondary in its relationship to the spoken norm, the main principle of the orthography of European (and many other) languages being the correspondence of a grapheme to a phoneme.

THE DESCRIPTION OF SPELLING-TO-SOUND RELATIONSHIPS IN ENGLISH, FRENCH AND RUSSIAN: PROGRESS, PROBLEMS AND PROSPECTS

Bruce L. Derwing, Tom M.S. Priestly, and Bernard L. Rochet

1. *Introduction*

This paper presents a summary of more than a decade of work on the problem of devising a systematic and reasonably simple framework for describing the spelling-sound (i.e., grapheme-phoneme) correspondences in the three major world languages indicated in the title. The initial effort was made by Derwing & Priestly for Russian, and resulted in a book (Derwing & Priestly 1980) that is now used in courses in phonology and phonetics. The application to English is being made by the same authors, while Rochet is working on a version dealing with French.

The application to Russian arose as a response to a specific problem that seemed to be endemic to students registered in 'Structure of Russian' courses: their marked inability to transcribe the language phonetically (even after several years of formal language study but, in the typical case, with little or no concomitant training in linguistics). Given the fact that these students were closely familiar with the standard orthographic representation of the language, this effort was soon directed towards the construction of a set of what the authors called *reading rules*, i.e., rules of "reading aloud", which took the standard spelling of any word sequence as input and served to derive a representation which is indicative of the way in which the string is ordinarily pronounced (see Derwing & Priestly 1980:4-5). The same approach is the basis for the applications to French and to English.

While the attempt to teach the pronunciation of a language through the spelling is by no means novel (in fact, spelling-based approaches are virtually universal, though often unsuccessful and misleading for students), we

do think that our efforts are unique in at least three important respects: (1) in trying to formulate a reasonably concise and non-redundant set of general ordered rules, we have achieved a level of *systematicity and homogeneity of description* which was lacking in other works along these lines;¹ (2) we have also tried to achieve, and (with respect to corpora of differing sizes in the three different languages, cf. 3.1. below) have achieved, a level of *comprehensiveness of coverage* and *attention to details* which were lacking in many previous works of this nature (e.g., Wijk 1966, Welna 1982, whose lists of examples often end with the word “etc.” and who appear to make insufficient attempt to weigh the alternative patterns on the basis of importance and/or frequency); and (3) we have given high priority to the matter of the *pedagogical utility* of our materials, with the goal of making them not only suitable for self-study, but also useful for learning other aspects of grammar and/or linguistic description beyond the details of pronunciation *per se*.

In accordance with this pedagogical aim, a number of decisions were made according to the criterion of *learnability* rather than on strictly theoretical grounds. So, for example, the rules for all three languages are *ordered*, not because we have any predisposition to prefer ordered rules, but because our rules happened to be easier to write, and easier to learn, under these conditions.

2. *Progress*

2.1. *Russian*

In line with this last-named goal, there quickly emerged in our description of Russian a distinction among three fundamentally different types of rules. The first of these we may call *spelling rules*, as they serve to compensate for ambiguities or inconsistencies in the orthography itself. Such rules operate on the standard orthography [SO] and convert it into a “regularized orthography” [RO], in which (ideally) all redundant characters are eliminated (e.g., the Russian ь, the “hard” or “separation” sign, and three of the so-called “soft” vowel-letters, viz., я, ë, ю) and inconsistencies in the spelling-system are reduced to a practical minimum (cf. the use of both ь and the “soft” vowel letters to indicate the palatalization of consonants on the one hand, and the use of both ÿ and the “soft” vowel letters to indicate the palatal fricative or glide sound [j]). Consider the following illustration, adapted from Derwing & Priestly (1980:17-22):

RULE 2Ai Я → ѧ / C— (where C = consonant letter)
 → ѧ / ELSEWHERE

These (ordered) rules have the following effects on the standard orthographic representations containing Я and/or ҃:²

SO	RO	
ПЯТЬ	п्यáть	(=/p'át'/)
ДЫНЯ	дýнъа	(=/dín'a/)
МОЙ	мойá	(=/mojá/)
ЯСНО	йáсно	(=/jásno/)
ПЬЯН	пýя́н	(=/p'ján/)
ИЗЪЯТЬ	изýáть	(=/izját'/)

The result is an RO in which not only the redundant letters Я and ҃ have been eliminated, but also one in which palatalization is consistently represented by the letter ѣ and the glide segment [j] by means of the letter ѣ.

As a direct consequence, there is a further advantage: viz., that the RO serves to indicate the morphemic structure of Russian words more clearly, in general, than does the SO, as illustrated below for four different case forms of some Russian nouns:

	SO	RO
N.Sg.	КНИГА	книг-а
A.Sg.	КНИГУ	книг-у
G.Pl.	КНИГ	книг-Ø
L.Pl.	КНИГАХ	книг-ах
N.Sg.	НЕДÉЛЯ	недéль-а
A.Sg.	НЕДÉЛЮ	недéль-у
G.Pl.	НЕДÉЛЬ	недéль-Ø
L.Pl.	НЕДÉЛЯХ	недéль-ах
N.Sg.	ÁРМИЯ	áрмий-а
A.Sg.	ÁРМИЮ	áрмий-у
G.Pl.	ÁРМИЙ	áрмий-Ø
L.Pl.	ÁРМИЯХ	áрмий-ах

Note that, in these examples, the N.Sg. ending is consistently represented as -a in the RO, the A.Sg. as -y, etc., a fact that is obscured in the SO representation of these forms.

The second main rule-set that has emerged from our work consists of

pronunciation rules, which are rules that describe how the standard (but now regularized) orthographic devices relate to a selected “phoneticized orthography” (PO), i.e., a type of representation in which (to the extent possible or practical) standard orthographic symbols are employed to indicate the actual pronunciation of the words. In standard Russian, for example, obstruents are assimilated in voicing to the last obstruent in the cluster (thus SO/RO ПРОБКА → PO пропка, SO/RO ПРОСЬБА → PO просьба) and are devoiced finally (SO/RO САД → PO сат, SO/RO ГВОЗДЬ → PO гвость), dentals are generally palatalized before other palatalized consonants (thus the form гвость just derived becomes the PO гвость, and SO ИЗЪЯТЬ → RO изять → PO изъять, etc.), while unstressed O is not distinguished from unstressed A, thus SO МОЯ → RO моя → PO майá and SO ЯСНО → RO ясно → PO ясна, etc.³ The resulting PO is a transcription that is roughly phonemic, i.e., it distinguishes all (and only) segments that contrast somewhere in the language and, as such, it can be converted to an equivalently broad “phonetic” transcription by a simple one-to-one transformation. Thus:

	Cyrillic		Latin
SO	RO	PO	PO
ПЯТЬ	пьять	пьять	→ /p'át'/
ДЫНЯ	дынья	дынья	→ /dín'a/
МОЯ	мойá	майá	→ /majá/
ЯСНО	йásно	йásна	→ /jásna/
ПЬЯН	пыйán	пыйáн	→ /p'ján/
ИЗЪЯТЬ	изъять	изъять	→ /iz'ját'/
ПРОБКА	прóбка	прóпка	→ /própká/
ПРОСЬБА	прóсьба	прóзьба	→ /próz'bá/
ГВОЗДЬ	гвóзьдь	гвосьть	→ /gvos't'/

Finally, such PO representations can be made more narrowly (and truly) phonetic by the subsequent application of a set of *phonetic detail rules*, such as the following ones for Russian:⁴

$$(1) \quad a \rightarrow \text{æ} / C' \leftarrow C'$$

(i.e., a stressed /a/ is fronted to [æ] between two palatalized consonants, thus PO /p'át'/ → [p'æt'], etc.)

$$(2) \quad a \rightarrow \text{ə} / \text{---} \text{ C}_0 \text{V} \\ \rightarrow \text{ə} / \text{---} \#$$

(i.e., unstressed /a/ is phonetically mid and tense in immediate pretonic syllables, but mid and lax at the end of words, e.g., PO /majá/ → [mʌjá], but PO /jásna/ → [jásnə]).

2.2. English

In the description for English (Derwing & Priestly, in progress), much the same kinds of distinctions among rule-types and among levels of representation have proven useful. (On the value of these concepts, see 2.4. below). For example, the following *spelling rules* can be used to eliminate the redundant letters c, q and x:⁵

- | | |
|---------|-------------------|
| RULE D | c → s / — {i,e,y} |
| | → k / ELSEWHERE |
| RULE H | qu → k / (1) — e# |
| | (2) LIST |
| | → kw / ELSEWHERE |
| RULE N1 | ex → egz / — V |
| RULE N2 | x → z / # — |
| | → ks / ELSEWHERE |

The effects of these rules may be summarized by means of the following examples:⁶

SO	RO
CITE	SITE
CELL	SELL
CY-cle	SY-kle
MOSQUE	MOSK
LIQU-or	LIK-or
QUIT	KWIT
SQUARE	SKWARE
ex-ACT	egz-AKT
XERX-es	ZERKS-es
SIX	SIKS
EX-it	EKS-it
OX-ide	OKS-ide
ex-CEED	eks-SEED

As an example of a pervasive inconsistency in standard English spelling, consider the use of the letter g to represent both the voiced velar stop

/g/ and the (contrastive) voiced palatal affricate /j/ which is often represented by the letter j). This inconsistency can be corrected by the following rule:

RULE E $g \rightarrow j / - \{i, e, y\}$.

This rule takes care of the “regular” cases such as SO GIN, AGE and ED-gy (respectively, RO JIN, AJE and ED-jy); but, as compared to RULE D which is almost exceptionless, there is quite a large number of exceptions to RULE E, which, for convenience, may be grouped into three separate classes:⁷

- (a) / g — (i.e., immediately after another g, as in BIGG-er and BUG-gy);
- (b) in words involving ng before the nominalizing suffix -er or the adjectivizing suffix -y (e.g., respectively, HANG-er, SING-er, and STRING-y, TANG-y);
- (c) in a long list of roots of Germanic origin, including such high-frequency words as GET, GIVE, GIRL, be-GIN, and 'HAM-BURG-er (see 3.2. below for a discussion of the question of regularity).

As for *pronunciation rules*, the most familiar English example is undoubtedly the rule that lengthens vowel nuclei in syllables containing a final “silent e.” Our complete formulation of this particular rule is as follows:

RULE DD $V \rightarrow \bar{V}$ 1. — $C_1^e \{#, -\}$
 $u \rightarrow \bar{u}$ 2. — $\{\#, -\}$

where C_1 = all consonantal letters and digraphs except {sh, ch, zh, j, ng}.

This rule lengthens all stressed vowels (whether primarily or secondarily stressed) — and all u’s, whether stressed or not — in two separate environments:

- (a) before a word- or syllable-final C_1^e (as in STAKE, DICE, STATE-ment, NICE-ly);
- (b) immediately before a word-boundary or a syllable-boundary (as in CHA-os, HE, DE-mon, LI-on, PO-et, GNU, RU-by, u-NITE).

Another, less familiar rule later affects these “long u’s”, as follows:

RULE HH $\bar{u} \rightarrow y\bar{o}\bar{o}$, except / C_d —, where $C_d = \{t, th, d, s, z, n, l, r, y\}$.⁸

This rule has the effect of inserting the glide element /y/ into words such as CUTE, a-BUSE, FU-ry, MU-sic, and so on, as well as in the unstressed syllables in, e.g., AM-bu-lance, AR-gu-ment, BI-'NOC-u-lars, u-NIQUE, etc.; this is why *all* u's, both stressed and unstressed, are included in rule DD above.

Finally (for our purposes here), English also exhibits the all-pervasive phenomenon of vowel-reduction, whereby most stressed vowels reduce to [ə], or (in some dialects) to [ɛ] or [ɪ], when unstressed (cf. 'TEL-e-GRAF and te-LEG-raph-y). Our formulation of this pronunciation rule is as follows:

RULE PP {a, ā, e, o, u, ia, ao, oa, oi} → ə.

Examples include OCE-an, (CAP-tain →) KAP-tān, MEL-o-dy, (FA-mous →) FA-maos, CUP-board, POR-poise. Compare the non-reduced vowels in AN, (TAINT →) TĀNT, LO, (MOUSE →) MAOSE, BOARD, POISE, and note that we choose the [ə]-dialectal pronunciation (cf. 3.1. below.)

As examples of low-level *phonetic detail* rules, we can mention the rule that lengthens vowels in all environments except before voiceless obstruents (cf. SEAT vs. SEED, SCENE, SEAL, SEA); the velarization of /l/ after vowels (cf. BALL, BALD vs. LEAF, LAUGH); the flapping of /t/ or /d/ between a stressed vowel and an unstressed vowel (in e.g. CIT-y, BOD-y), and the aspiration of voiceless stops in a variety of environments, e.g. at the beginning of a word (cf. PIT vs. SPIT).

2.3. French

In the description for French (Rochet, in progress), the same three-way division among type of rule may also be made (see 2.4. below).

Because of the common Romance lexicon shared by French and English, a number of *spelling rules* look very similar in the two languages. Thus, English rule D has for its equivalent French rule 10:

RULE 10 c → s / — {i, e, y}
 → k / ELSEWHERE
 e.g., ceci, cycle; cape, copier.

Rule 10, however, must be accompanied by a French-specific rule:

RULE 15 ç → s

Similarly, the French counterpart for English rule E is the following:

RULE 20 $g \rightarrow j / — \{i, e, y\}$
 e.g., nager, agir, gymnase.

Rule 20 must be supplemented by the following:

RULE 25 $ge \rightarrow j / — V (V) (C)$
 e.g., pigeon, mangeaille.

French rule 30 is also similar to English rule H but it differs from it in its scope and in the number of exceptions it admits:

RULE 30 $qu \rightarrow k$
 (except in adéquat, équateur (where $qu \rightarrow kw$), and in piqûre (where $qû \rightarrow kü$))

Rule 40 deals with the voiced counterpart of *qu*:

RULE 40 $gu \rightarrow g$
 except in linguiste and its derivatives (where $gu \rightarrow gq$) and lingual and its derivatives (where $gu \rightarrow gw$)⁹.

Among French *pronunciation rules*, the most common are those dealing with the semi-vowels (50-58), and the mid vowel adjustment rules (60-68), for example:

RULE 50 $i \rightarrow y / — V$
 e.g., pied, chantier, mariage.

RULE 54 $u \rightarrow w / — V$
 e.g., chouette, couenne.

RULE 58 $\ddot{u} \rightarrow \ddot{w} / — V$
 e.g., fuite, écuelle.¹⁰

RULE 60 $O \rightarrow [o] / — \#$
 $\rightarrow [ɔ] / — C \{ \$, \# \}$ ¹¹

RULE 64 $Œ \rightarrow [ö] / — \pm$
 $\rightarrow [\alpha] / — C \{ \$, \# \}$
 e.g., peu, vœu; peur, cœur.

RULE 68 $E \rightarrow [e] / — \#$
 $\rightarrow [\varepsilon] / — C \{ \$, \# \}$ ¹¹
 e.g., passer, parlez; fer, amer.

French *phonetic detail rules* include the lengthening of stressed vowels by a following voiced fricative, and the lengthening of nasal vowels in stressed checked syllables:

RULE 70 $V \rightarrow V: / — \{v, z, r, \check{z}\} \#$
 e.g., louve, chose, dure, tige.

RULE 74 $\check{V} \rightarrow \check{V}: / — C \#$
 e.g., blanche, mince.

One may also mention here the specification of unstressed mid-vowels in free syllables using two alternate approaches, viz., either on the basis of dialectal and stylistic characteristics:

RULE 69₁ $E \rightarrow [e] / — \$ (C) V (+ \text{stress}) \{ \text{informal} \}$
 $\rightarrow [\epsilon] — \$ (C) V (+ \text{stress}) \{ \text{formal} \}$

or in terms of vowel harmony:

RULE 69₂ $E \rightarrow [e] / — (C) V (+ \text{stress}, - \text{low})$
 e.g., laisser.
 $\rightarrow [\epsilon] / — (C) V (+ \text{stress}, + \text{low})$
 e.g., aigreur, dessert.

2.4. On "Regularization"

The distinction between *spelling rules* and *pronunciation rules*, made for the sets of spelling-to-sound derivations for all three languages concerned, is embodied in the transcription derived by the first rule-type, the *REGULARIZED ORTHOGRAPHY (RO)*, upon which the rules of the second type operate.¹²

It was emphasized in 2.1. that the RO for Russian not only eliminates the totally redundant letters of the Standard alphabet, and renders the representation of consonant-palatalization very much more consistent, but also indicates the morphemic structure of Russian words far more clearly than is shown in the Standard Orthography.

In 2.2., we stated that the distinction between spelling-rules and pronunciation-rules for English was "useful"; and the same may be said for the distinction for French. As far as the rules for these languages are concerned, we consider it *at least* expedient to derive an RO in which only the recognized letters of the Standard Orthography are used. This RO is characterized, as we see it, by *minimal ambiguity*, viz., it can be read aloud with at least an acceptable pronunciation and without difficulty; and by *minimal unconventionality*, i.e., it comes as close as possible (within the limits otherwise imposed) to normal spellings — e.g., for English, we re-spell the SO BREAK-fast as BREK-fast, to ensure the correct pronunciation of the first syllable, but balk at other logical possibilities, such as

BREKK-fast or BREK-fust, which are unnecessarily unconventional.

Neither the RO for English nor the RO for French has the *morphemic transparency* of the RO for Russian, and to this extent they are therefore less valuable. Moreover, the allocation of certain of our rules (for either language) to either the spelling rules or the pronunciation rules is made pragmatically: for example, a possibly counter-intuitive allocation of a particular rule to one or the other subset is chosen if the formulation of that rule is found to be thereby much simpler. To this extent, pedagogical convenience outweighs any *a priori* theoretical considerations based upon the actual form of the rules concerned. We do not claim, however, that we have determined a pedagogically perfect formulation for all of our rules, and the potential clash between theory and practice in the selection of the “perfect” RO may eventually not be all that jarring.

The rules by which our English RO is derived do, in some instances, contradict considerations of *frequency*; e.g., rule D eliminates the letter c, which is far more common than the letter k in the SO. Similarly, rule E changes the common g to the uncommon j. We do not rule out the possibility that such choices may result in something pedagogically unsound: viz., “derived” spellings that may instil poor spelling-habits in students. This possibility must be borne in mind, and psycholinguistic testing procedures will probably prove valuable.

3. *Problems*

3.1. *Preliminary Decisions*

Although by this time our description of Russian is complete and the analyses of English and French nearly so, it is of some interest to consider here at least a few of the major problems we encountered along the way, including one or two whose resolution still remains somewhat unclear.

Obviously, a number of critical preliminary decisions had to be made even before any project of this type could get under way, notably (1) the particular (dia-)lect(s) to be described, (2) the level of (phonetic) detail to be included, and (3) the total amount of (lexical and grammatical) coverage to be achieved. Our decisions with respect to these particular questions varied somewhat as a function of the language being described.

In the Russian case, where standard or “academic” norms have been explicitly and officially prescribed, there was good reason to select these, at least as the primary focus of attention. This meant the “Literary Standard”, described in greatest detail by Avanesov (see Derwing & Priestly 1980:4, 8-

9; Avanesov 1972).

For French, we follow for the most part the description given by Lerond (1980), which, although essentially a description of Standard or Parisian French, grants to vowel neutralization the place it deserves and thus allows more readily for dialectal and stylistic variation, and for recent shifts or phonetic changes in progress.

For English, where no such prescriptive standards have been widely accepted, there was much more room to maneuver (and much more uncertainty); and in one respect our work has parted company with most of the prior sources. For purely pragmatic reasons we have chosen to break with the tradition that has selected RP British English (as a highly-prized prestige norm), preferring to opt for a widespread North American dialect instead. The rationale for this is that most non-native students who want or need to learn English nowadays tend to choose North American rather than British English, on current social, political and economic grounds. As for the choice of particular dialect features, we have generally operated within the limits specified by Webster (1970), Morris (1979), or Kenyon and Knott (1953). Where controversial cases have arisen, we have resolved these in each instance by choosing that dialectal variety which would result in the simplest form of the rule in question (and, where appropriate, the shortest list of exceptions). One important specific consequence of this decision was to choose a dialect norm which, within the allowable limits, neutralized as many distinctions as possible. Thus, for example, whereas in the dialect of Derwing (the one native North American among the three authors) a systematic distinction obtains between most words in "short o" (e.g., COT, SOT, DON, FOG, JOB, CLOCK), which are pronounced with the low back unrounded [ɑ]¹³, and a parallel set of "aw", "au" and "ou" words (e.g., CAUGHT, SOUGHT, DAWN, DRAW, COUGH, etc.), which have the rounded [ɔ], it was considered simpler to opt for the neutralizing case where all these words are pronounced in the same way, and all the spellings are converted to a single orthographic shape.¹⁴ As a practical matter, the chosen digraph, *aw*, may be realized in two different phonetic ways: either as [ɑ], which occurs in a neutralizing dialect spoken in much of California, or as [ɔ], which occurs in the neutralizing dialect spoken throughout most of Canada. At this low level, of course, the user of our materials is quite free to make his or her own choice: both dialects are widespread, and neither introduces any major problems of intelligibility.

With respect to *phonetic detail*, we agree again quite pragmatically in our choices for all three languages. Because these materials were intended

primarily as a pedagogical aid for non-native learners of the languages involved, a pronunciation that maintained all of the potentially contrasting (i.e. traditional *phonemic*) distinctions was the best that we could realistically hope for, and also one that would certainly suffice for purposes of ordinary face-to-face communication. With this as our stated goal, further “phonetic refinement” was likely to be of interest only to native speakers seeking a more thorough analytic treatment of their own language.

Finally, with respect to *extent of coverage*, we were somewhat more ambitious in our rules for Russian than in those we wrote for the other two languages. Given both (a) the notable regularity of the Cyrillic orthography in its application to Russian, and (b) the detailed treatment of exceptions provided in Avanesov & Ožegov (1960), we tried to make our treatment of Russian *absolutely complete*; i.e., to list in our “Lexicon of Exceptions” *all* of the counter-examples that had come to our attention. For English, faced with the notoriously inconsistent character of its orthographic norms, we lowered our sights to the achievement of the following specific goal: to account for all forms whose Standard Frequency Index was 36.0 or higher as determined by Carroll, Richman & Davies (1971). According to this source, this amounts, roughly, to the first 25,000 most frequent English words, and includes every word that appears with an average frequency of 1 (or more than 1) in every 1,000,000 words of running text.¹⁵ Even to achieve this level of completeness for English, while still producing lists of exceptions that were of manageable size, we had to make the following omissions:

1. all numbers and other non-standard orthographic symbols (e.g., \$, %, &);
2. all abbreviations and acronyms;
3. the names of the letters of the alphabet;
4. all *regular* noun-compounds (i.e., those whose pronunciation can be predicted from the pronunciation of their component parts);
5. all *regular* inflections (e.g., plural and possessive forms of nouns; progressive, past participle, preterite and 3rd person singular present forms of verbs; comparatives and superlatives of adjectives);
6. all *regular*¹⁶ examples of the following highly frequent classes of derived forms (which could also be treated in a similar way to noun-compounds): deverbal nouns with the suffixes -er, -ment; denominal adjectives with the suffixes -less, -able, -ful, -ive, -y; and deadjectival adverbs with the suffix -ly.

For French, we adopted as a starting point the first 5,082 words of Juillard (1970).¹⁷ We then combined these 5,082 words with the 5,000 word list compiled by Matoré (1963) according to his “conscience linguistique”, by eliminating from larger dictionaries the words he considered less useful. This procedure was considered necessary in view of our pedagogical objectives, because Juillard’s list did not include some concrete words such as *fourchette*, which must be considered an essential part of the vocabulary available to native speakers and which should be available to second language students.¹⁸ Combining Juillard’s and Matoré’s lists yielded a total of 6,751 words.

3.2. *Problems of Description*

The remaining major problems involve specifics of description. In some instances, these vary dramatically as a function of the language involved. As already mentioned, the status and treatment of *stress* and *syllable-division* are two such problem areas.

For Russian, stress can not be assigned systematically. We therefore advise students of Russian simply to memorize the placement of stress on every new form, or else to resort to dictionary look-up. In this way, stress is treated as part of the SO representation; it is taken as a “given” feature of the spelling of every Russian word whose pronunciation is derived by our rules (see Derwing & Priestly 1980:11).

For English, although serious attempts to account for/predict the position of stress have been made, with varying degrees of success (see, e.g., Chomsky & Halle 1968:59-162 and an attempted application by Schnitzer 1979), all are far too complex to satisfy our criterion of practical utilization by non-native language-learners. We therefore treat stress in much the same way for English as for Russian, though with a different notation.¹⁹ Faced, moreover, with what we call the *Demon-Lemon Problem* (or, alternatively, the *Evil-Devil*), it is in general the case that, in English words of the type exemplified by these slogans, one can predict the length/tenseness of the vowels from the position of the syllable-boundaries, *or* the reverse; but one cannot predict both. Of the two courses of action, we chose the former, and thus show the syllable-divisions in “demon” and “lemon” in our lexicon (thus: DE-mon, LEM-on) and derive vowel-length/tenseness therefrom. We thus indicate *both* stress and syllabification for *all* words within our range of responsibility (hence the need for the systematic exclusion mentioned above; cf. Wełna 1982:14-16). The contrast between DE-mon and LEM-on (or between E-vil and DEV-il) contains the necessary

information for the student to apply rules DD and PP above.²⁰

For French, stress always falls on the last pronounced syllable of a word, and presents no problems. Syllabification is also very straightforward. So as to be most effective, syllabification rules are applied twice in French, once to the standard orthography (the rules given at that level can be justified on independent grounds in that they are the end-of-line hyphenation rules), and once to some form of phoneticized representation. Both sets of rules are very similar, the second being slightly more detailed than the first (which will be henceforth referred to as hyphenation rules). The cover symbols used here are V = any vowel, C = any consonant, O = obstruent, L = liquid, and G = glide.

<u>Hyphenation rules</u>	<u>Syllabification rules</u>
H1: V C V → V - C V	S1: V V → V \$ V
H2: V O L V → V - O L V	S2: V C V → V \$ C V
H3: V C C V → V C - C V	S3: V O L V → V \$ O L V
	S4: V C G V → V \$ C G V
	S5: V C C V → V C \$ C V

Hyphenation is useful for determining the value of the letter e according to the following rule:

RULE 33 e → ə / — { -, # }
 → E / — C { -, # }

(i.e., e represents schwa when it occurs in a free or “open” orthographic syllable; it represents a mid-front unrounded vowel when it occurs in a checked or “closed” syllable), e.g.,

	RECHERCHER	DESSERT
H	re-cher-cher	des-sert
33	rə-chEr-chEr	dEs-sErt

Syllabification rules are then applied to the output of pronunciation rules (and in particular those which simplify consonant sequences and delete final consonants, labelled here 37 and 38, respectively):

	rəchErchEr	dEs\$Ert
37	rəšEršER	dEsErt
38	rəšEršE	dEsEr
S	rə\$šEr\$šE	dE\$šEr

At that point, a phonetic value may be assigned to the abstract rep-

resentation E (or to that used for the other mid-vowels), on the basis of the type of syllable in which this representation occurs, following rules 68 and 69₁:

	rə\$\$Er\$\$E	dE\$\$sEr
68	rə\$\$s̄er\$\$s̄e	dE\$\$s̄er ²¹
69 ₁	rə\$\$s̄er\$\$s̄e	de\$\$s̄er

For the purposes of syllabification, therefore, no elaborate lexical lists are required for French.

Even with these stipulations, however, a number of important general descriptive problems remain in the treatments of all three languages, and especially for English and French. These involve, in particular, the following:

1. the problem of distinguishing the rule from the exceptional case, and
2. the problem of assessing just how many examples are required to justify (even pragmatically) the proposal of a rule in the first place.

The following English example serves to illustrate the first of these problems. Even after we have excluded all the words from Carroll, Richman & Davies (1971) below the critical frequency figure (cf. above), and all the forms designated by the supplementary criteria (cf. above), our lexicon still yields quite a large number of words in which the orthographic sequence "ea" represents both [i] (e.g., LEAF, BEAD, STEAL) and [ɛ] (e.g., BREAD, DEAF, THREAT). Which of these sets constitutes the regularity, and which the exception? Judging by total *token* frequencies of the two sets of words in Carroll, Richman & Davies (1971), the pattern "ea" = [i] wins easily; using *type* frequency the result is the same, but the winning margin is much narrower. On the assumption that these frequency counts are reliable, we therefore take the "ea" = [i] pattern to be the regular case. We would be hard-pressed to make our decision if either (1) the token and the type frequencies were in conflict, or (2) some other word count were to suggest a different result. Perhaps the best test would be psycholinguistic in character, involving, e.g., the giving of a set of novel (nonsense) words to (literate) native speakers in order to determine which of the pronunciations is actually preferred.

Data from English also illustrate the second major descriptive problem noted above, viz., trying to decide whether a limited number of examples is sufficient to justify the positing of a viable spelling or pronunciation rule.

Another of our proposed rules for English is as follows:

RULE LL $a \rightarrow o / w$ — *EXCEPT BEFORE {k, g, ng, w}*

(i.e., stressed a (= [æ]) becomes o (= [ɔ]) after w, with the exceptions noted.) The only examples on our list that utilize this rule are *nine* in number: WHAT (→ WAT), WAR, WARM, WANT, WATCH, SWAN, WAN-der, QUAR-rel (→ KWAR-rel), and QUAL-i-ty (→ KWAL-i-ty); on the other hand, words with the exceptional consonants listed number *six*; QUACK (→ KWAKK), WAX (→ WAKS), WAG-on, TWANG, SQUAW (→ SKWAW), SQUAWK (→ SKWAWK). The ratio of nine examples to six exceptions may well not justify the particular formulation that we have chosen for our rule; but once again, the crucial question appears to be one of frequency, *and* what the native speaker “thinks” of the matter. Indications of possible tests of the validity of our formulation come from (a) rare words such as “quagga” QUAG-ga (→ KWAG-ga), with a pronunciation [kwægə], rather than [kwɔgə], as is predicated by our rule; and (b) the disparate pronunciations that native English speakers tend to give to the *same* Chinese family name in its two different spellings, “Wong” and “Wang” — the former reflecting the Chinese pronunciation [wɔŋ] and the latter the Anglicized [wæŋ].

3.3. Grammatical Information

It is of interest to compare our spelling-to-sound rules for the three different languages from another perspective: namely, the extent to which we have cited *grammatical information* in the formulation of our rules. Again, we have recourse to formulations of this nature for purely pragmatic reasons; but these comparisons do indicate something about the three orthographic systems.

For Russian, first, many of our rules depend on specific boundary-symbols, which in turn are defined (in part) with reference to grammatical entities (namely, prefixes and prepositions). Apart from this, there are no fewer than *seven* separate rules which cite grammatical categories: 3C1 applies to the genitive singulars of adjectives; 3G3 and 3L refer to reflexive verbs; 3M, to participial forms; 3T, to verbal stems; and both 3V and 3X, to specific case-forms of nouns and adjectives. All these instances of recourse to grammatical information parallel the complexities of the pronunciations prescribed for Contemporary Standard Russian (see the relevant sections in Derwing & Priestly 1980 for details).

For English, only *one* rule specifies grammatical information directly:

RR relates to the pronunciation of “s” and “d” in inflectional endings. Elsewhere, admittedly, grammatical information is referred to -- both in our Lexicon, e.g., the noun CON-duct as opposed to the verb con-DUCT, and in the discussion of some of the rules, e.g., the noun USE as opposed to the verb USE, only the latter of which has to undergo the change of s to z (cf. Venezky 1970:41-44, Cronnell 1971:3-4).

For French, a number of rules are sensitive to grammatical information. Thus the sequence -ent in word-final position represents schwa if it is the 3rd person plural verb marker, or the nasal vowel [ã] in all other cases, as indicated in the following ordered rules:

RULE 22 ent → ə / — [3rd person verb marker]

RULE 23 en → [ã] / — C #

In this case, grammatical information is absolutely essential in order to obtain the correct output.

In order to determine whether word-final r preceded by the letter e is pronounced or not, one must consider both phonetic (or graphemic) and grammatical characteristics of the string:

RULE 34 r → Ø / e — #

(a) when the sequence er is preceded by a palatal ([y], [š], [ž]), i.e., in the endings -ier, -yer, -iller, -cher and -ger, except in monosyllables, e.g., papier, broyer, poulailleur, boucher, berger but fier, hier, cher.

(b) in the infinitive of 1st conjugation verbs, e.g., passer, attraper, manquer, peler.

In all other cases,

RULE 35 r → [r] / e — #

e.g., amer, éther, hiver, fer.

It could be argued that recourse to grammatical information should be interpreted as an indication of failure for anyone attempting to predict the phonetic values corresponding to orthographic strings. However, this failure is only real when one is dealing with automated translation (from spelling to sound).²³ In second language teaching, on the other hand, one does not expect reading ability to be developed independently and to the exclusion of all other linguistic skills. Students should therefore be capable of drawing — and encouraged to draw — on their knowledge of the language as a whole in order to decipher its orthographic code.²⁴

4. *Prospects*

To save space, the prospects that we foresee are presented here very concisely.

4.1. *Testing*

First, we have started to test our rules for English and French by computer. This ties in with automation, see 4.3. below.

Second, we plan to test the psycholinguistic validity of our rules. Such possibilities have already been mentioned (cf. the reference to “quagga” above.) Another example: we can check on the validity of the rule which derives ‘yōō’ from u in open syllables in word-initial position by testing subjects with rare words such as “ubrant” and “usufruct” (note that the rule may not be a good one for English, given the number of borrowings with u- which are pronounced without the ‘y’, such as “upas”, “ulu”.)

4.2. *The “Ideal RO”*

In 2.4 we discuss the usefulness and validity of our Regularized Orthography. However elusive this may be, we consider it worth pursuing, not only on theoretical grounds, but for its practical utility in discussions of spelling reform and language spelling.

4.3. *Other Applications*

Although summarized here only briefly, the following potential applications are perhaps the most important:

The teaching of reading and the remediation of reading and/or pronunciation problems;

Automated translation from spelling to phonetic representation, and thence to sound, as in reading machines;

Psycholinguistics, e.g., in the testing of selected reading theories, and in the spelling of nonsense words with particular phonetic properties for various testing situations.

NOTES

1. Cf. Avanesov & Ožegov 1960 for Russian; Hill & Ure 1962 or Wijk 1966 for (British) English; Cronnell 1971 for (American) English. These consist essentially of long, non-integrated — and often overlapping and redundant — lists of graphemes and grapheme sequences, together with all of their possible “pronunciations”. See also Venezky 1970, Wełna 1982 and

Börner 1977, whose rules are not fully integrated, but are much more systematic and less redundant than those in the other works cited.

2. Some phonetic details are specified by subsequent rules, as explained below.
3. Our rules predict the place of stress for neither Russian nor English; for these two languages, this information must be looked up in a dictionary (cf. 3.2.). Note the use of CAPITAL LETTERS to show the representation in the Standard Orthography for Russian.
4. Our rules also include adjustments for some of the (non-“standard”) styles typical of more natural, rapid or “allegro” speech, but none of these are discussed here (see Derwing & Priestly 1980:130-179 for details).
5. The rule-labels, while arbitrary, are indicative of priorities in ordering.
6. For English, CAPITAL LETTERS are used to indicate the syllables that are stressed (“China” and “china” being both represented as CHI-na), and the hyphen to indicate a syllable-boundary. Problems arising are discussed in 3 below.
7. GENERAL exceptions can be categorized in three sets: (1) idiosyncratic and seldom used (e.g., “forecastle”, “victuals”); (2) idiosyncratic and occurring in common words (e.g., “does”, “handkerchief”); and (3) systematic and occurring in uncommon words “o” in “covet”) or in common words (“o” in “love”). Of these, we omit (1) entirely; we list (2) in the lexicon; and we both list in the lexicon, and devote special sections to, (3). Exceptions to SPECIFIC RULES can be categorized in the same three kinds of groups, e.g., with reference to the pronunciation of *c* and *g*, (1) “Celtic” (in its scientific, not its sporting, use); (2) “soccer”; and (3) the set of ‘Germanic-origin’ words given immediately below.
8. In some dialects, of course, rule HH applies after some of these consonants also: cf. RP British /tyōōn/ for “tune”, South Wales English /blyōō/ for “blue” (see Wells 1982:206-208). Problems of (dia-)lectal variation are discussed in 3.
9. Note that RULE 40 must follow RULE 20 if the pronunciation of *gui* ([gi]) is to be derived correctly.
10. Here some phonotactic restrictions could be handled by means of subrules, e.g., $i \rightarrow [y] / -V$ except in the environment / O L —, where O = obstruent and L = liquid (e.g., *crier*, *prier*, *plier*), or through ordered rules. Thus, rule 50 could be applied in its generality and the incorrect outputs *krye*, *prye* and *plye* could be adjusted by means of a subsequent phonetic detail rule (see below), such as $y \rightarrow i / O L —$. This possible double treatment points to the fact that the three categories of rules described here are by no means absolute and hermetic. The same rule may appear under different headings depending on the analysis adopted.
11. The symbol \$ is used in the rules for French to represent a syllable boundary. For the value of the symbol E, see below, rule 33.
12. For theoretical discussion of rule-types, see Venezky 1970:45-46, 123; 1976:23-24; Cronnell 1971:3. See also note 10.
13. There are a very few unusual exceptions, e.g. DOG [dɒg] and COF-fee [kɔfi], see Wells 1982:136-137.
14. Note also the decision not to show any contrast between “merry”, “Mary” and “marry”, and cf. Cronnell 1971:7-8.
15. Venezky 1970 deals with a “20,000 word vocabulary”.

16. For all three of these *regular* categories of omissions, the problem, of course, is not one of listing exceptions (since, by definition, such forms may not constitute exceptions), but rather one of specifying which vowels/syllables are *stressed* (for both Russian and English) and where syllable-*boundaries* occur (for English). For this purpose, a special auxiliary set of rules is required.

17. Although Juilland aimed at a list of the 5,000 most frequent words, he included all the words with a combined coefficient (for frequency and dispersion of usage) of 3.00 or more. This procedure isolated a basic vocabulary of 5,082 words; see Juilland (1970: LXXIII).

18. Juilland (1970: XXVII) uses the term "vocabulaire disponible" and points out that the same situation arose when *le français fondamental* was produced.

19. For Russian, stress is indicated with diacritics (the acute and grave accents for primarily- and secondarily stressed vowels) alone. For English, primarily- and secondarily stressed syllables are shown in capitals; the distinction between the two is indicated by using the apostrophe for the former (e.g., CON-quer; de-SCEND; CON-de-SCEND). Note that *syllables* must also be specified for English, since the orthography involves the use of vowels with no syllabic value: compare the first *e* in "lateness" (LATE-ness) with the first *e* in "lateral" (LAT-er-al). Cf. Cronnell 1971:3.

20. We thus end up with a system much like that used in many well-known dictionaries; these tend to be unsystematic, however (cf. Webster's representation of "division" and "divisor" as *di-vision* and *di-vi-sor*, showing that grammatical information has intruded on the phonological criterion!)

21. Lerond (1980) leaves all unstressed mid-vowels unspecified to allow for variation. Thus he represents the words *rechercher* and *dessert* as *rə\$šEr\$še* and *dE\$ser*. While retaining E in the first syllable of *dessert* is an acceptable solution — allowing both [dE\$ser] and [de\$ser] — the retention of E in the second syllable of *rechercher* must be rejected because it allows the non-attested pronunciation [rəšerše], whereas only [rəšerše] is attested. Lerond's position must therefore be modified to state that mid-vowels should be specified only in word-final syllables and in checked unstressed syllables.

22. On regularity and the productivity of rules, see Venezky 1970:39-41 and especially 123-129; Venezky 1976:22-23; Cronnell 1971:4; and Glushko 1981, especially 64-65.

23. Even then, ways can be found to specify strings which will identify the item in question as belonging or not belonging to a given category.

24. Although an elaboration of this point is beyond the scope of this paper, it is worth mentioning that formulating spelling-to-sound rules often reveals interesting insights into various aspects of the structure of the language considered, and in particular, its morphophonemics.

BIBLIOGRAPHY

Avanesov, R.I. 1972. *Russkoe literaturnoe proiznošenie*. 5th ed. Moscow:

Prosvětanie.

Avanesov, R.I. & S.I. Ožegov. 1960. *Russkoe literaturnoe proiznošenie*.

- Slovar'-spravočnik.* 3rd ed. Moscow: Goz. izd. inostrannyx i náacional'nyx slovarej.
- Börner, Wolfgang. 1977. *Die französische Orthographie.* Tübingen: Max Niemeyer Verlag.
- Carroll, John B., Barry M. Richman & Peter Davies. 1971. *The American heritage word frequency book.* Boston: Houghton Mifflin.
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English.* New York: Harper & Row.
- Cronnells, Bruce. 1971. *Annotated spelling to sound correspondence rules.* Southwest Regional Laboratory, Technical Report 32.
- Derwing, Bruce L. & Tom M.S. Priestly. 1980. *Reading rules for Russian. A systematic approach to Russian spelling and pronunciation.* Columbus, Ohio: Slavica.
- . (in progress). *From letter to sound in American English: Learning the pronunciation from the spelling.*
- Dewey, G. 1970. *Relative frequency of English spellings.* New York: Teachers College Press.
- Glushko, Robert J. 1981. "Principles for pronouncing print: the psychology of phonography." in: A.M. Lesgold & C.A. Perfetti (Eds.) *Interactive processes in reading.* Hillsdale, N.J.: Erlbaum, 61-84.
- Hill, L.A. & J.M. Ure. 1962. *English sounds and spellings.* London: Oxford University Press.
- Juillard, Alphonse. 1970. *Frequency dictionary of French words.* The Hague/Paris: Mouton.
- Kenyon, John S. & Thomas A. Knott. 1953. *A pronouncing dictionary of American English.* Springfield, Mass.: Merriam.
- Lerond, Alain. 1980. *Dictionnaire de la prononciation.* Paris: Larousse.
- Martinon, Philippe. 1913. *Comment on prononce le français.* Paris: Larousse.
- Matoré, Georges. 1963. *Dictionnaire du vocabulaire essentiel: les 5000 mots fondamentaux.* Paris: Larousse.
- Monnot, Michel. 1971. "Examen comparatif des tendances de syllabation dans les mots abrégés de l'anglais et du français." *Le français moderne* 39: 191-206.
- Morris, William (Ed.) 1979. *The American heritage dictionary of the English language.* Boston: Houghton Mifflin.
- Peretz, Caroline. 1977. "Aspects sociolinguistiques du parler parisien con-

- temporain.” In H. Walter (Ed.) *Phonologie et société*, Montréal/Paris/Bruxelles.
- Richman, Marie. 1976. “The Distribution of /e/-/ɛ/ in French verb endings,” *Phonetica* 33: 307-319.
- Rochet, Bernard L. (in progress). *Reading rules for French* (title uncertain).
- Schnitzer, Marc L. 1979. “Generative orthography,” Unpublished MS.
- Venezky, Richard L. 1970. *The structure of English orthography*. The Hague: Mouton.
- . 1976. *Theoretical and experimental base for teaching reading*. The Hague: Mouton.
- Warnant, Léon. *Dictionnaire de la prononciation française*. Deuxième édition, revue et corrigée. Gembloux: Duculot.
- Webster. 1970. *Webster's seventh new collegiate dictionary*. Toronto: Allen.
- Wells, John C. 1982. *Accents of English*. Cambridge: Cambridge University Press.
- Wełna, Jerzy. 1982. *English spelling and pronunciation*. Warsaw: Państwowe wydawnictwo naukowe.
- Wijk, Axel. 1966. *Rules for the pronunciation of the English language*. Oxford: Oxford University Press.

THE AUTOMATED PHONETIC TRANSCRIPTION OF ENGLISH TEXT

S.G.C. Lawrence, B.J. Williams, and G. Kaye

1.0 *Abstract*

There have been many reports of rule systems for automated text-to-speech transcription. These have, with varying degrees of success, produced, as an intermediate stage, a phonemic transcription of the input. This paper describes a rule system that takes an input text stream in conventional orthography and produces

1. A phonemic transcription
2. An allophonic transcription
3. A phonetic transcription with some consonantal assimilation.

At present the speaker model is RP, but other accents are equally possible within the general system design.

The system has been implemented on an IBM Personal Computer. It runs in real time and displays results in conventional IPA with diacritics, thus giving the user immediate feedback.

Examples illustrating the quality of transcriptions will be given.

2.0 *Introduction*

The work reported here is part of a larger project to produce, from machine-readable text, acceptable and intelligible synthetic speech which can be listened to for extended periods without it causing fatigue.

A typical use of such a system is in the correction and proofing of documents. Lee (1957) has shown that the ear is more efficient at detecting mistakes than is the eye. However, for this to be exploited it is essential that any irregularity in the pronunciation of the synthetic speech be the result of

an error in the input text, rather than an artifact of the system.

For this objective to be realized an accurate phonetic transcription must be generated from the input text. This transcription should be based on a well-defined and acceptable model of English pronunciation (RP has been chosen in the present work). It should represent how the text might be read by a competent and careful reader; for example, a BBC news reader. A system has been developed to meet this objective. Its three stages produce:

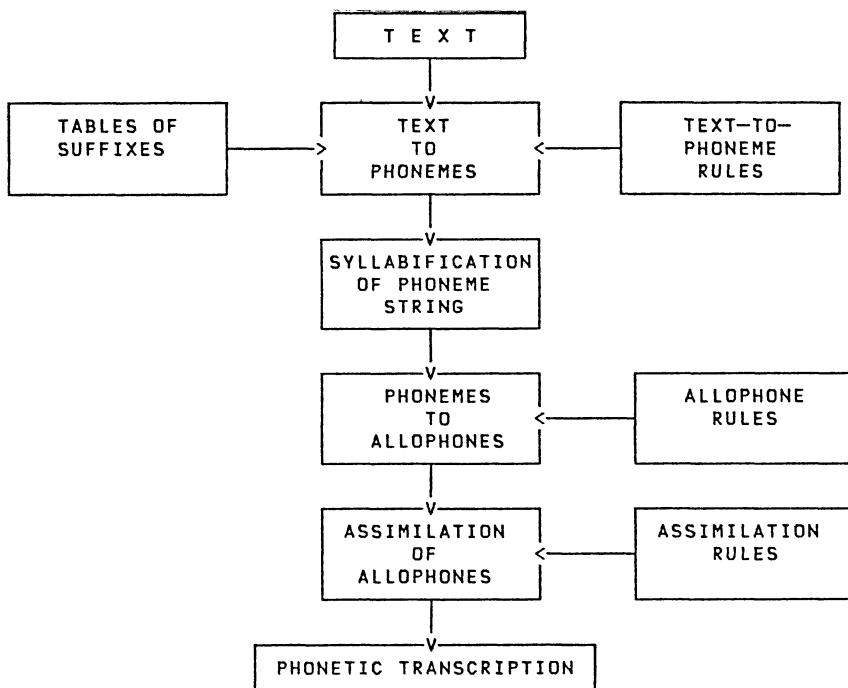
1. A phonemic transcription
2. An allophonic transcription
3. A phonetic transcription with some consonantal assimilation

Each stage is driven by a set of rules stored apart from the program code. The present version is flexible in its design and implementation. Thus experimental changes may be made to the rules, and the effect checked by the user. It has been developed to run on an IBM Personal Computer with a modified character set that includes the International Phonetic Alphabet and the more important diacritics.

This paper gives a general discussion of the three rule systems, highlighting the methodology and potential weakness. It is illustrated by a sample transcription.

3.0 *Overall View*

The figure below shows the stages involved in the transcription. The text is preprocessed to convert all numbers to words and abbreviations are analysed to determine whether they are to be spelled out letter-by-letter (eg. 'UN') or pronounced as a word (eg. 'UNICEF'). This system takes no account of lexical stress. At present, all vowel reduction is embodied within the transcription rules. Such an approach has obvious pitfalls, leading at times to incorrect vowel reduction, and lack of compactness of the rules. A new system has been developed exploiting the techniques of lexical stress assignment (Williams 1986). This system does not require the text-to-phoneme rules to account for vowel reduction, which is conditioned by the lexical stress rules. Another restriction is that no attempt has been made to resolve the pronunciation of homographs. For example, the pronunciation of the word 'house' depends on whether it is a noun (/haus/) or a verb (/hauz/). This requires the assignment of syntactic categories which is to be handled by the new version of the system.



4.0 *The Rules*

Many rule schemes for transcribing text into phonemes have been published, notably those of Ainsworth (1973), McIlroy (1974), and Elovitz et al. (1976). These all aim to produce some form of phonemic transcription. None of them produce an allophonic transcription, nor do they take account of consonantal assimilation. The present work follows a strategy similar to that of Elovitz, with extensive modifications and additions to handle the allophonic transcription and the assimilations.

The rules for each of the three stages have the same basic format:

$$\text{LEFT-C(object)RIGHT-C} = \text{transcription}$$

The string being processed is scanned left-to-right. A match is sought between one or more characters in the input string and the 'object' section of the rules. When such a match is found the left and right contexts (LEFT-C and RIGHT-C) are checked in turn. If all three match then the contents of 'transcription' are added to the output string. The search continues with

the following characters in the input string until the whole of the input has been transcribed.

4.1 *Phonemic Transcription Rules*

The text string is scanned from left to right in a single pass to produce phonemes from the graphemes using the set of rules. The rules are grouped alphabetically by the initial letter of the ‘object’ string and listed in a critical order with a default last rule for each letter, which acts as a catch-all.

For example, the rule

$$i (a) b, l = \emptyset$$

produces the schwa in the word ‘variable’ /veəriəbl/.

If a match with a rule fails, the next rule is tried. Repeated failures lead to the last rule for the given letter, which always produces a successful match. The last rule for the latter ‘a’ is

$$(a) = \emptyset$$

As can be seen by the last ‘a’ rule the RIGHT-C and LEFT-C can be null. The transcription string may also be null. For example, when two identical consonants occur in a word they are pronounced as one phoneme. The rule for omitting the second ‘m’ in ‘summer’ is

$$m (m) =$$

The LEFT-C and RIGHT-C strings may contain sets of letters, and tests are made for exclusion from the membership of the set, as in one of the ‘e’ rules:

$$[p, s, v] (e) [t, p] = \emptyset$$

This rule would be successful for ‘reset’ but not for ‘lever’. It should be pointed out that although this rule would apply to ‘compete’ there are earlier rules which would also apply to this word, thus transcribing the second syllable of it before the above rule is encountered. Therefore the ordering of the rules is of critical importance.

The facility for specifying that an item in the input string must be either a member of a set [a, b, c, d] or not a member of set [x, y, z], makes for compact specification of the rules and improves processing efficiency.

A novel feature of this rule schema is the ability to match the LEFT-C string with the phonemic transcription already produced. When ‘ed’ occurs at the end of a word it is pronounced ‘t’ when preceded by a voiceless

sound. The appropriate rule is:

[/p, k, f, θ, s, /ʃ/] (ed) # = t

(The ‘#’ is used as a symbol for end-of-word.)

Functions can also be specified in the RIGHT-C and LEFT-C strings to increase the generality of the rules, i.e.:

- COM Find zero or more consonants.
- C1 Find one consonant.
- C1M Find one or more consonants.
- C2I Find two identical consonants.
- SBX Find a voiceless sibilant (i.e. one of s, /ʃ/). This function can only occur in the LEFT-C string, as it refers to the phonemes already generated.
- SFOM Find zero or more suffixes.
- SF1M Find one or more suffixes.
- SFV1 Find one vowel suffix. This function is similar to SF1M, except that the first suffix must begin with a vowel for the function to be successful. This allows implementation of the ‘magic-e’ rule where the ‘e’ has been lost in suffixation.
- SIB Find a sibilant (i.e. one of s, /ʃ/, z, ʒ, tʃ, dʒ). This function can only occur in the LEFT-C string, as it refers to the phonemes already generated.
- V1 Find one vowel.
- V1M Find one or more vowels.
- VF1 Find a front vowel (i.e. e, i, y).

4.1.1 Example of Text-to-phoneme Rules

The following is an example of some of the rules for the letter ‘e’.

e001	- [a], CIM(e)g, e	= i
e004	(easy) #	= i:, z, ɪ
e009	(eque)	= e, k
e010	f, o, r, g(et)	= e, t
e011	#, g(et)	= e, t
e012	/k, g, tʃ, dʒ/(et)COM, #	= ɪ, t
e013	C1, C1(et)COM, #	= e, t
e037	/z, tʃ, ʃ, ʒ, dʒ, s/](es) #	= ɪ, z
e038	- [/#/], [/p, t, k, f, θ/](es) #	= s
e039	V1, COM(es) #	= z

e040	(e)f, u, l, SFOM, #	=
e041	#, V1M(e)C1M, #	= e
e088	V1, COM, [t, d](ed) #	= i, d
e089	[/p, k, f, θ, s, /ʃ](ed) #	= t
e090	V1, COM(ed) #	= d
e107	V1, C1M(e)SFOM, #	=
e108	V1(e) #	=
e109	#, C1M(e)SFOM, #	= i:
e110	(eter)SFOM, #	= i, t, ə, *
e111	(e)C1, e	= i:
e112	(e)C1, SFV1, s, #	= i:
e113	(e)C1, i	= i:
e114	(e)n, C1	= e
e115	(e)C1, SFV1, #	= i:
e117	(e)	= i

Notes

- ‘*’ in the ‘transcription’ string signifies a potential linking – ‘r’ (see rule ‘e110’). After processing the text, the string of phonemes is scanned for ‘*’ and is replaced by ‘r’ when the following sound is a vowel.
- Complete words may be contained in the rules.
- The identification code on the left is used in the statistical analysis of the operation of the rules.
- The symbol ‘–’ is the negative operator, applied to a grapheme or phoneme set (see rule ‘e001’).

4.2 *The Allophonic Transcription and Assimilation Rules*

The rules for allophone production and assimilation are similar in format to the phonemic transcription rules, but contain different functions, relating to features of phonemes rather than graphemes. The ‘object’ and ‘transcription’ strings may also contain functions. One sound may belong to several classes of features: for example ‘t’ is voiceless, plosive, aveolar, obstruent, and consonantal. Each phoneme in the ‘object’ is processed by all the rules. There are some phonemes for which no allophone or assimilation rules exist.

The following function mnemonics are used only in the allophone and assimilation rules.

–VOX	A voiceless sound – p,t,k,f,θ,s,ʃ,h
ALVEOLAR	An alveolar – t,d,l,s,z,n,r
APPROX	An approximant – w,r,l,j

CONSONANT	A consonant – p,b,t,d,k,g,m,n,l,r,f,v,s,z,h,w,tʃ,dʒ,ŋ, θ,ð,ʃ,ʒ,j,ɿ,l,
DENTAL	A dental – θ,ð
FRIC+VOX	A voiced fricative – v,ð,z,ʒ
FRONTVOWEL	A front vowel – ε,i,eɪ
GEMINATE	A second instance of the preceding phoneme
LATERAL	A lateral – l
NASAL	A nasal – n,m,ŋ
NASAL@X	A nasal at X, where X is the place of articulation
OBSTRUENT	An obstruent – p,b,t,d,k,g,s,f,v,z,tʃ,dʒ,θ,ð,ʃ,ʒ
ROUND	A bilabial consonant or rounded vowel – p,b,m,w,u,ʊ,ɔ,ɒ
STOP	A stop – p,t,k,b,d,g
STOP+VOX	A voiced stop – b,d,g
STOP–VOX	A voiceless stop – p,t,k
STOP@X	A stop at X, where X is the place of articulation
VELAR	A velar – ɳ,k,g
VOWEL	A vowel – ɑ,æ,ɛ,ə,i,ɪ,ɒ,ɔ,u,ʊ,ʌ,ɔɪ,au,əʊ,ɛə,ʊə,ɪə,ai,ei,ɔɪ,ɔə,ɜ
VOWEL+STRESS	A stressed vowel
VOWEL–STRESS	An unstressed vowel
VOX	A voiced sound i.e. not –VOX
NOVOWEL	Find no succeeding vowel in the syllable
VSTRESS	Find a stressed vowel in the same syllable

The following is a list of allophonic properties or ‘attributes’ used in the ‘transcription’ string of the allophonic rules.

ASPIRATED
DENTALISED
DEVOICED
GLOTTALISED
LABIALISED
LENGTH (length increases from 1 to 6)
NASALISED
PALATALISED
PART—ASPIRATED
PART—VOICED
UNEXPLODED

VELARISED

4.2.1 *Example of Allophonic Rules*

At the level of the phoneme-to-allophone rules, an attempt has been made to encode linguistically-motivated rules for the modification of phonemes in context. These rules are a development of the articulatory-based distinctive features of Ladefoged (1975), and provide a modelling of the articulatory dimension for what is essentially a terminal-analogue formant synthesiser.

The following is an example of some of the allophone rules.

1	(VOWEL)	FRIC+VOX	= LENGTH,6
2	(VOWEL)	STOP+VOX	= LENGTH,5
3	(VOWEL)	—	= LENGTH,4
4	(VOWEL)	[LATERAL, NASAL]	= LENGTH,3
5	(VOWEL)	FRIC-VOX	= LENGTH,2
6	(VOWEL)	STOP-VOX	= LENGTH,1
7	(STOP)	[—], STOP	= UNEXPLODED
8	— (STOP-VOX)	—[APPROX], VSTRESS	= ASPIRATED
9	-[s] (STOP-VOX)	VOWEL-STRESS	= PART-ASPIRATED

Notes

1. The symbol ‘—’ is used to represent a syllable boundary or a word boundary (see rule 3). The output from the text-to-phoneme rules is divided into syllables, and syllable boundary markers are inserted into the string of phonemes.
2. The symbol ‘|’ is used to indicate that the following set is optional.
3. Square brackets (‘[]’) are used to delimit a set, of which only one member may occur at the given position.

4.2.2 *Example of Assimilation Rules*

The following is an example of some of the assimilation rules. The set of rules has been partly based on the work of Wells and Colson (1971).

1	(n)	[STOP@X,NASAL X]	= NASAL@X
2	(STOP)	STOP@X,VOWEL	= STOP@X
3	(STOP)	NASAL@X	= STOP@X

4.3 *Sample Output from the System*

The following is an example of the output produced by the transcription system.

ORTHOGRAPHY

The very first showroom is used by young mums with handbags

BROAD PHONEMIC TRANSCRIPTION

ðə verɪ fɜːst ʃəvrum ɪz juːzd baɪ jʌŋ mʌmz wið hændbægz

SYLLABIFICATION

ðə ve-ri fɜːst ʃəv-ru:m ɪz ju:zd baɪ jʌŋ mʌmz wið hænd-bægz

ALLOPHONIC TRANSCRIPTION

ðəverifɜːst²ʃəvu:mɪzjuːzd⁰bajl~ŋml~mzw iðhæ~ndbægz.
4 4 4 24 4 6 6 4 6 5

ALLOPHONE ASSIMILATION

ðəverifɜːʃʃəvu:mɪzuːzbajl~ŋml~mzw iðhæ~mbægz.
4 4 4 2 4 6 6 4 6 5

Notes

1. The diacritics follow the phoneme to which they apply.
2. The symbol ‘0’ indicates glottalisation of a stop.
3. The numbers underneath give a length value for certain phonemes. The length ranges from 1 to 6, the shortest being 1. Normal length (3) is unmarked notationally.

4.4 Checking the Rules

A dictionary containing solely the head-words with their pronunciation and part of speech is being abstracted from Collins' English dictionary. It will be used to tag with their phonemic transcription all the words occurring in the LOB corpus (1978), which contains the frequency of occurrence of individual words in 1,000,000 words of written text published in Britain in 1961. There are approximately 25,000 words that occur more than once. This tagged corpus will be used to check the accuracy of the rules against the most frequently-used words in English. Statistics can be gathered to show which rules have been used for a given word, and whether the correct pronunciation was produced. So far the text-to-phoneme rules have been checked against a corpus of 1800 frequently-used words, the score obtained being more than 90% success. Currently there are more than 950 rules. This number should decrease when rules are deleted in the light of the insights gained by checking against the LOB corpus, and when the modified approach described below is implemented.

5.0 *Improvements*

The rule system presented above has been shown to give good transcriptions for a large part of the English vocabulary. However there are weaknesses in the system. The basic strategy is to execute the transcription as a monotonic left-to-right process, so that once a sequence of graphemes has been transcribed the resulting transcription cannot subsequently be modified. Such an approach means that the known phonotactic constraints of English cannot be called upon explicitly to filter out non-occurring phoneme sequences according to their position in the word or syllable. However, a variable length of preceding and following context can be taken into account when transcribing a sequence. This allows some indirect application of phonotactic constraints by means of their representation in the orthography. Spellings such as 'knee', 'gnat', 'bomb' and 'sing' can be handled in this way. Alternatively, the transcriptions could proceed as follows.

knee	gnat	bomb	sing
{kni:}	{gnæt}	{bɒm}	{sɪŋg}
/nɪ:/	/næt/	/bɒm/	/sɪŋ/

where the non-occurring consonant clusters are removed by the application of phonotactic rules. Cases such as the above can be handled by the orthography with reasonable ease. A different situation arises in the cases where /æ/ becomes /ɒ/ when preceded by /w/ and followed by any consonant except one of /k/, /g/ or /ŋ/. This rule would account for cases such as 'wad' /wɒd/ and 'whack' /wæk/, and would need to 'look at' both the left and right environments in the phonemic rather than the orthographic form. This entails a more complex processing strategy involving backtracking, in order to take account of the right environment in phonemic form, so that the modified strategy is not a pure left-to-right monotonic process. The 'magic-e' rule is ordered before this rule, to account for the diphthong in words such as 'wade' and 'wake'.

A more extensive use of the currently-generated phoneme string as the LEFT-C should also improve the compactness and accuracy of the rules. However, a different scanning strategy must be followed if phonotactic constraints are to be more directly applied. In particular, the strategy will have to allow backtracking in the transcription process.

The present strategy attempts to carry out vowel reduction without any

reference to lexical stress. This is one of the more common causes of errors. A modified set of rules is being developed which will in most cases not attempt vowel reduction, but will reserve this to a later stage, according to the strategy outlined in Williams (1986). The example below uses words beginning in ‘theo-’ to demonstrate that any strategy which does not take account of lexical stress is almost bound to fail on at least some words. The transcription produced by the rules is shown, together with the modified transcription which is produced after the application of rules of vowel reduction conditioned by lexical stress. The first example word has primary (lexical) stress on ‘the-’, the second example has secondary stress on ‘the-’, and the third example has unstressed ‘the-’.

orthography	preliminary phonemic transcription	phonemic transcription with lexical stress and consequent vowel reduction
theorem	/θi:ɒr+em/	/'θiərəm/
theological	/θi:ɒ+lɒdʒɪkl/	/,θiə'lɒdʒɪkl/
theology	/θi:+ɒlɒdʒi/	/θi 'ɒlɒdʒi/

6.0 Acknowledgements

It is a pleasure to thank William Collins Sons and Co. for making available for our use a copy of the Collins English Dictionary in computer-readable form.

BIBLIOGRAPHY

- Ainsworth, W.A. 1973. “A system for converting English text into speech,” *IEEE transactions on audio and electroacoustics*, Vol. AU-21, No. 3, June 1973.
- Elovitz, H.S., R.W. Johnson, S. McHugh, J.E. Shore. 1976. “Automatic translation of English text to phonetics by means of letter-to-sound rules,” U.S. Naval Research Laboratory Report 7948, January 1976.
- Johnson, P.W. and C.J. Lovell. 1980. “Universal contextual translator,” *IBM Technical Disclosure Bulletin*, Vol. 23, No. 1, June 1980.
- Ladefoged, P. 1975. *A course in phonetics*. New York: Harcourt Brace Jovanovich, Inc.

- Lee, W.R. 1957. "Spelling irregularity and reading difficulty in English." Slough: National Foundation for Educational Research. Occasional Paper No. 2.
- LOB (Lancaster-Oslo/Bergen) Corpus. 1978. Department of English, University of Oslo.
- McIlroy, M.D. 1974. *Synthetic English speech by rule*. Computing Science Technical Report #14, Bell Laboratories, March 1974 (revised September 1974).
- Wells, J.C. and G. Colson. 1971. *Practical phonetics*. London: Pitman.
- Williams, B.J. 1986. "English word stress in a text-to-speech synthesis system". IBM UKSC Report no. 136.

PHONOLOGICAL ACCESS TO DATA BASES

L. Hitzenberger

1. *The access problem for alphabetic files.*

Phonology is a very theoretical domain. In connection with computer science there is a chance for it to become an applied science, part of which is the access problem to files.

Access to traditional alphabetic files like telephone books, catalogues, etc., usually implies exact knowledge of the key element, mostly the name of the item. But the spelling of names is especially variable, because normalized spellings (according to the development of writing) exist alongside antiquated forms which do not correspond to modern spelling. (The spelling *Th* in *Thor* or *Thür* in German was abolished during the 19th century, but it is still frequent in proper names.)

One possibility of resolving this problem is to normalize spelling. This is the way taken by certain library catalogue instructions (e.g., the Prussian Instructions of 1899) which prescribe certain spellings and thus change the proper form of names. Thus, they do not distinguish between *Müller* with *u* umlaut and with *u, e*.

This way of resolving the problem cannot be chosen for most modern applications of name files, because normally users are not ready to learn the rules for the changes and there are legal reasons for the binding form of names. .

With the use of electronic data processing there arise new problems in the access to name files. On a printed medium you have a certain overview because of the arrangement in pages. This often allows finding the requested item even if the exact position is not known, because users then tend to browse through the pages. This effect has to be simulated in automatic access, because otherwise the success rate of the edp-medium would lie behind that of the printed medium. This means that it is necessary to intro-

duce a certain vagueness in automatic access to name files in data bases.

Generally there are four indices for names:

- (1.) the concrete orthographic string
- (2.) the phonological string
- (3.) any code (for instance a personal number) or a hash code*
- (4.) a hash code with a phonologically based algorithm, which maintains the natural connection between the name and the hash code.

Numbers 2 and 4 are the domain of automatic phonology.

2. *Automatic Phonology*

Automatic phonological procedures consist of an algorithm representing the rules, or an algorithm to process a phonological grammar (cf. Hitzenberger 1980 for a description of PANAMA, which is a program system to process a phonological grammar of the type proposed in Chomsky & Halle 1968).

2.1 *Phonological search in name files.*

There are some well-known algorithms which use the phonological and orthographical properties of language to produce a code which reduces the variety of orthographic signs to a smaller number of symbols. The Soundex-Côdes consist in their simplest form of a very restricted grammar, which may look approximately like the following:

Table 1: Soundex-Grammar

$\left\{ \begin{matrix} B \\ P \\ V \\ F \end{matrix} \right\} \rightarrow 1$	$\left\{ \begin{matrix} C \\ G \\ J \\ K \\ Q \\ S \\ Z \end{matrix} \right\} \rightarrow 2$	$\left\{ \begin{matrix} A \\ E \\ I \\ O \\ U \\ H \\ W \\ Y \end{matrix} \right\} \rightarrow 0$
$\left\{ \begin{matrix} D \\ T \end{matrix} \right\} \rightarrow 3$		
L → 4	$\left\{ \begin{matrix} M \\ N \end{matrix} \right\} \rightarrow 5$	R → 6

* A hash code is any index, where the key element is generated according to some fixed algorithm.

These procedures are simple heuristics and combine two principles of reduction:

1. phonetic equivalence
2. phonological-orthographical similarities.

2.1 Phonetic equivalence

Phonetic equivalence means that because of the ambiguity of the letter-sound-relation there is more than one orthographic string for one phonetic string. The frequency of these ambiguities in German is not very high and they are mainly restricted to the vowels (e.g. I, Y; E, Ä; EI, ÄI; ...) and to word endings (*Auslautverhärtung*).

The above mentioned codes like the Soundex-Codes also have as a result the reduction of the phonetic equivalences, but this is just because of a much further-going reduction, which immediately fails, if you need a more precise phonological chain. So, by eliminating the vowels it is of course possible to find all variants of the name *Meier*, but it is not possible to indicate that *Maier* and *Meier* have the same phonetic representation, but not *Meer*, *Moor* or *Mohair*. Also it is not possible to establish that *Moor* and *Maar*, for instance, are more closely related to each other than *Moor* and *Mir*.

2.2 Phonological-orthographical similarities

These similarities are important for any retrieval-system on alphabetic data. If one does not know the correct orthographic string of a name, there is a high probability that there is a regularity between its recollection and the salient features of the proper name. This regularity is partly based on phonological properties. But there may also be syntactic, semantic or other influences. Some of the phonological properties are considered in the Soundex-Codes by building classes. But these classes do not express the phonological interdependencies and are static. There are no means to express that /m/ belongs to the class of nasals as well as bilabials. I will return to this point later on.

The result is that sign-oriented algorithms are not adequate, and that you have to use the theory of distinctive features to resolve the problem.

The application of a well-founded theory does not imply that we have to integrate all implications of this theory. It is the genuine domain of computer and information science to select those parts of general language theories which are important for man-machine communication and to adapt them for a functional solution. This also means that we do not have to

look for complete solutions but for effective ones for man-machine communication. We are searching for optimums but not for maximums.

Thus, for the application of phonology to the access problem it is possible to use a very restricted phonology with an inventory of maximally eight features and a very restricted rule corpus.

3. *Design of a generative phonology for data-base access.*

At the University of Regensburg we are developing an algorithm for phonological access to a telephone information system.

The general idea of the system is to translate the orthographic string of a name into a fairly narrow phonological representation and to generate from this form in turn all possible orthographic representations again.

The basis for the analysis of a corpus of 10,000 various names was the phonological parser PANAMA (Hitzenberger 1980). This parser is able to process a phonological grammar of the *SPE*-type (Chomsky & Halle, 1968) including all variables, scalar features, boundaries and so on.

The “Regensburger Phonology” (Hitzenberger 1986) uses the following system of six features:

Table 2: Features: Vowels

	a	o	u	ä	e	ö	ü	i
M1	+	+	+	+	+	+	+	+
M2	—	—	—	—	—	—	—	—
M3	+	+	+	—	—	—	—	—
M4	—	+	+	—	—	+	+	+
M5	+	—	—	+	—	—	—	+
M6	—	+	—	—	—	+	—	—

The features are named M1 to M6, because the phonological correspondence is not always unique. Thus, M3 is the feature “dark” for the vowels, but it has an ambiguous meaning for the consonants. The reason for the distribution of the features is a fairly simple heuristic which aims at getting phonologically meaningful classes by eliminating higher rank features. But the features should by no means describe the phonemes in a

Table 3: Features: Consonants

	g	k	b	p	d	t	v	f	s	ʃ	x	m	n	r	l	ts	h
M1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+
M2	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	+
M3	-	-	-	-	-	-	+	+	+	+	+	+	+	-	-	+	-
M4	-	-	+	+	+	+	+	+	-	-	-	-	-	-	-	+	-
M5	+	+	+	+	-	-	-	-	+	+	-	-	-	+	-	+	-
M6	+	-	+	-	+	-	+	-	+	-	+	-	-	-	-	-	+

phonologically correct sense. The letters of the orthographical strings are translated into feature complexes according to context-free rules:

Table 4: Context-free Rules

A	/a/	I, J	/i/	R	/r/
B	/b/	K	/k/	S	/s/
C	/ts/	L	/l/	T	/t/
D	/d/	M	/m/	U	/u/
E	/e/	N	/n/	V	/v/
F	/f/	O	/o/	W	/v/
G	/g/	P	/p/	X	/ks/
H	/h/	Q	/k/	Y	/i/
				Z	/s/

To this representation apply the context-dependent rules:

Table 5: Context-dependent Rules

1. Simplify digraphs
2. H → Ø
3. Vowels:

ei	→	ai	ae	→	ä
eu	→	oi	oe	→	ö
ie	→	i	ue	→	ü
äu	→	oi			

4. Consonants:	$\text{sch} \rightarrow \int$
	$\text{ph} \rightarrow \text{f}$
	$\text{ch} \rightarrow \left\{ \begin{array}{l} \text{k} / \left\{ \begin{array}{l} \# \\ \text{s} \end{array} \right\} \\ \text{x}_c \end{array} \right\}$
	$\text{c} \rightarrow \left\{ \begin{array}{l} \text{k} / — \\ \text{—} / \left\{ \begin{array}{l} \text{a} \\ \text{o} \\ \text{u} \\ \text{Kons} \end{array} \right\} \\ \text{—} / \left\{ \begin{array}{l} \emptyset \\ \text{k} \end{array} \right\} \\ \text{—} / \text{s} \end{array} \right\}$
	$\text{v} \rightarrow \text{f} / — \#$

With these rules we achieve a representation of very high-level phonology in which orthographical ambiguities are reduced.

4. *Extension of the access possibilities*

To extend the possibilities of access beyond the phonetic representation, we can now reduce the number of the features, and thus reduce the number of phonemes. Because of the hierarchical order of the features we can eliminate them one by one and build bigger and bigger classes. The way to form the classes can immediately be seen by considering the binary trees corresponding to the vowels and the consonants.

Table 6: *Binary Tree for Vowels*

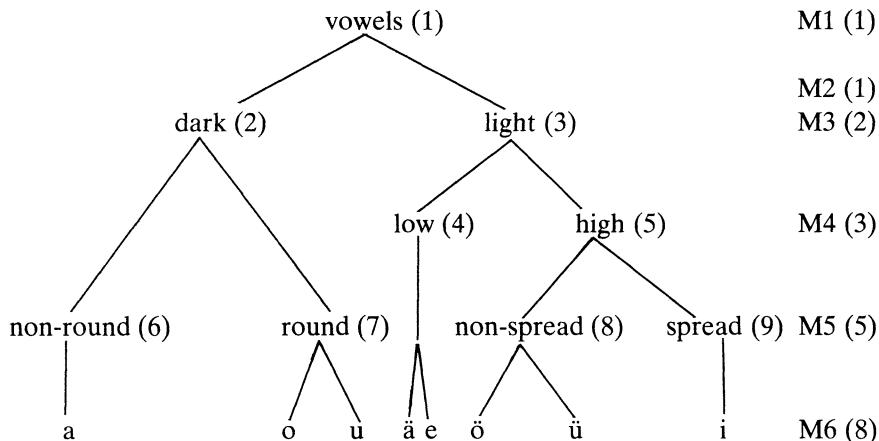
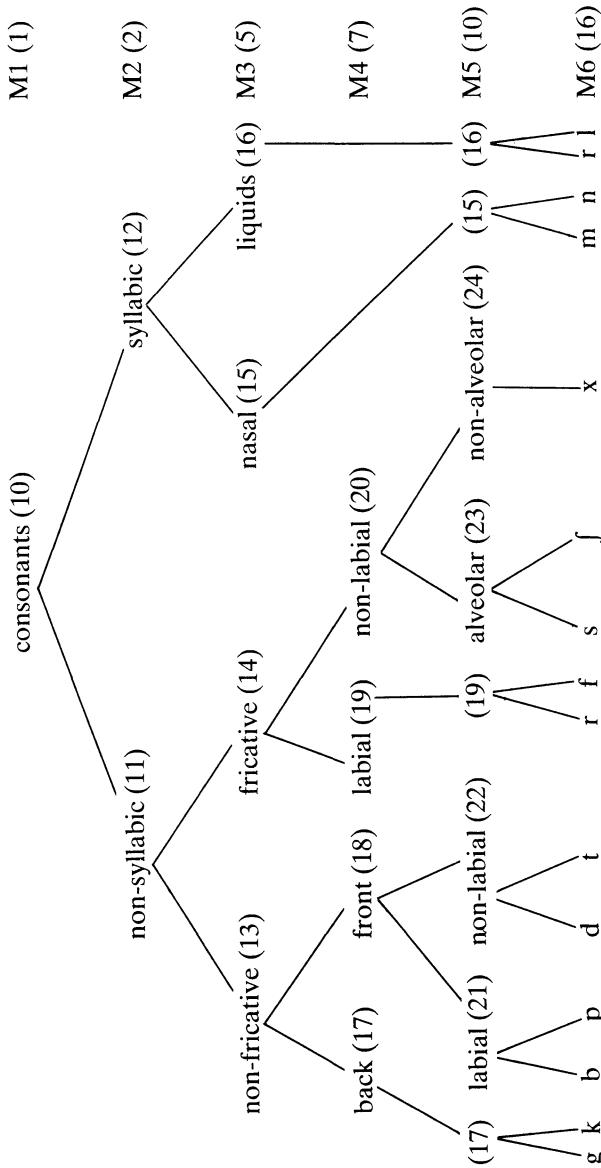


Table 7: Binary Tree for Consonants



The numbers after the names of the features indicate how many different representations we have on the various levels. Thus we achieve a dynamic process of coping with phonological similarity.

The example in Table 8 shows the gradual collapsing of the various representations into one.

Table 8: Orthographic levels of representation

chain	M6	M5	M4	M3
AAHOEH	aö	6,8	2,5	2,3
AHU	au	6,7	2,4	2,3
ARTEL	artel	<i>6,r,22,4,I</i>	2,16,18,4,16	2,16,13,3,16
ARDELE	ardel	<i>6,r,22,4,I</i>	2,16,18,4,16	2,16,13,3,16
ARBELL	arbel	<i>6,r,21,4,1</i>	2,16,18,4,16	2,16,13,3,16
ALKOHOL	alkol	6,1,17,7,1	2,16,17,2,16	2,16,13,2,16

(The numbers correspond to the nodes of the trees. The italicized elements show the results of a simple Soundex-Code like the Kölner Phonetik (Postel 1969)).

The theoretically possible level of M2 or M1 shows practically nothing but the vowel-consonant cluster of the word.

5. *The experimental application of a variety of grammars.*

One of the goals of the Regensburger Phonology was the generation of an optimal hash-code of five bytes length. There should be an optimal distribution of the attachment of approximately five to fifteen names to one key element. Within this one key element, we use different procedures to identify the proper entry. To achieve this distribution we applied a variety of grammars to our corpus of names, with various weights on the importance of the vowels and the consonants.

To summarize the results, we can say that in general phonological grammars give a very good basis for code generation. So the ratio of unique codes in a five-letter code with the grammar of Table 5 and six features is 65%, with five features it is 53% and with four features 26%. The typical distribution of the keys is shown in Table 9.

Table 9

Number of names attached to one key	Number of keys for 6 features	for 5 features	for 4 features
1	6538	5355	2683
2	888	1056	947
3	175	280	395
4	92	118	242
5	24	64	131
6	20	26	71
7	14	24	59
8	11	11	33
9	4	3	23
10	4	8	24
11	2	7	18
12	0	3	6
13	0	0	7
14	1	1	10
15	0	0	8
16	0	0	5
17	0	0	3
18	0	0	1
19	0	0	2
21	1	1	1
29	1	1	1
34	1	1	1
160	1	1	0
162	0	0	1

The reduction of vowels does not effect the distribution very much, because of the low functional load of the vowels. This means that we can easily reduce the applied grammar to a grammar of consonants and delete the vowels.

Moreover, to come to a further reduction of the variety of keys we can also reduce the number of consonants by joining the plosives and some of

the fricatives. So we revise our grammar of Table 5 by changing the rule for the vowels to

$$\text{vow} \rightarrow \emptyset$$

and adding

$$\begin{bmatrix} - M4 \\ - M5 \\ - M6 \end{bmatrix} \rightarrow k$$

$$\begin{bmatrix} + M2 \\ + M3 \\ + M4 \end{bmatrix} \rightarrow x_c$$

at the end of the consonant rules. This grammar brings us the fairly optimal distribution shown in Table 10:

Table 10

Number of names attached to one key	Number of keys for 6 features	for 5 features	for 4 features
1	5174	3968	1303
2	953	940	609
3	278	355	342
4	149	184	189
5	72	84	131
6	39	61	82
7	21	39	59
8	18	24	43
9	7	16	49
10	8	11	28
11	5	10	24
12	1	9	24
13	2	8	20
14	2	2	15
15	1	5	13
16	1	1	9

17	2	3	9
18	0	2	8
19	0	1	4
20	1	0	3
21	0	2	1
22	1	1	3
23	0	0	1
24	0	0	1
25	0	0	1
26	0	0	8
27	0	0	2
28	0	0	1
29	1	0	0
30	0	1	2
31	0	0	2
32	0	0	3
33	0	0	1
34	1	1	1
36	0	0	1
37	0	0	2
40	0	0	1
42	0	0	1
43	0	0	1
44	0	0	1
45	0	0	1
47	0	0	1
160	1	1	0
162	0	0	1

BIBLIOGRAPHY

- Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
- Hitzenberger, Ludwig. 1980. *Automatisierung und Phonologie: Automatisierte generative Phonologie am Beispiel des Französischen*. Linguistische Arbeiten 91. Tübingen: Max Niemeyer Verlag.
- Hitzenberger, Ludwig. 1986. *Möglichkeiten bei der Anwendung*

- phonologischer Verfahren bei der Namenssuche in Datenbanken*, In:
Peter Hellwig (Ed.) *Trends in der linguistischen Datenverarbeitung.*
Jahrestagung der GLDV 1984. Hildesheim: Olms.
- Instruktionen für die alphabetischen Kataloge der preußischen Bibliotheken*
vom 10. Mai 1899 (Preußische Instruktionen). Reprint. Wiesbaden: Otto
Harrassowitz, 1964.
- Postel, Hans Joachim. 1969. *Die Kölner Phonetik*. *IBM-Nachrichten* 19, p.
928.

ON LINGUISTIC ERROR*

Philip A. Luelsdorff

0. *Introduction*

Our interest in linguistic errors stems from the commonly held view (cf. Fromkin 1973, 1980) that errors provide clues to the underlying linguistic knowledge of the linguistic sender-receiver (speaker-writer-speller-hearer-reader), the processing of that knowledge in language comprehension and production (Garret 1981), and that error analysis subserves the evaluation of linguistic argumentation (Stemberger 1983). In the cases of both normal and pathological linguistic data, errors are seen as evidence which may support or refute theories of the constraints on the range of possible mispragmatic, missyntactic, missemantic, misphonological and misphonetic (cf. Maxwell 1982 for a theory of constraints on pathological misarticulation phonologies), and misspelling (cf. Luelsdorff 1986) systems, and the typological investigation of such systems. In the present paper we present an extended view of the notion ‘linguistic error’ within the organon-model of language, a discussion of the notion ‘possible linguistic error’, a critique of an established notion of linguistic sign and revision thereof designed to predict the major occurring error types.

1. *The notion ‘linguistic error’.*

In this section we sketch an extended theory of the notion ‘linguistic error’, extended in the sense that it is grounded in the ever-expanding organon-model of language, a theory of factors and functions ultimately derived from Plato’s *Cratylus*, introduced into linguistic theory by Bühler (1920, 1931, 1934, 1936, 1969) and developed by Jakobson (1960), Hymes (1961,

* This paper is a revision of section I.3 of Luelsdorff 1986.

1968), Habermas (1973, 1976a, 1976b, 1976c, 1976d), Holenstein (1979), and Luelsdorff (1984). The organon-model of language is a model of *language qua speaker's competence* (communicative competence) as manifested by speech events in context, what we have been terming "differential linguistics" (Luelsdorff 1982). This theory of linguistic factors and functions contains the notion of a norm-theoretical speech event in L_i at time T_i (both psychological and sociological). From the conjunction of templates (cf. 3. below) and cooccurrence restrictions which constitute the linguistic norm, the sender-receiver, within the confines of his competence, selects, combines and recombines just those conceptual archetypes which conform, as a rule, to those semantically representable attitudes and intentions he wishes to actualize (foreground) by using his linguistic performance as an instrument to communicate them. Note that we distinguish between the class of such possible speech events in L_i , on the one hand, and the class of those events which a particular sender-receiver of speech events in L_i is capable of producing/comprehending and actually produces and comprehends, on the other. Clearly, there is a vast difference between the norms of a language and the norms of language which an individual sender-receiver has internalized, on the one hand, and what an individual sender-receiver is capable of producing or intends to produce as opposed to what he actually produces, on the other. Very clear evidence of the matter is provided by linguistic errors, although there are, of course, linguistic errors which are intentionally produced and intentionally received as such, suggesting the necessity of the notion 'differential error-function'.

For detailed accounts of the factors and functions of speech events which have thusfar been discovered, we refer the interested reader to the discussions cited above, restricting ourselves here for lack of space to the following summary list:

<i>Factor</i>	<i>Function</i>	<i>Validity Claim</i>
Speaker	Expressive	Sincerity
Hearer	Appellative	Normativity
Objects	Referential	Adequacy
Channel	Phatic	Equality
Code	Metalinguistic	Intelligibility
Message	Aesthetic	Beauty
Utterance	Communicative	Interpersonal Subjectivity
Inner Form	Constitutive	Executability

Attitude	Adaptive	Conformity
Intention	Selective	Clarity

Each of the above categories in the theory of differential linguistics has the status of a variable realized by various values. Each configuration of extramicrolinguistic values may have microlinguistic reflexes in the items selected from and sequenced in the particular linguistic code for the particular linguistic utterance, which choice and combination occupies a place in a system of multidimensional coordinates constituting the conventions commonly referred to by the rubric "norm". The individual sender-receiver may or may not subscribe to the norm in question, either intentionally or unintentionally, either in knowledge or in ignorance. While norms themselves do contain variants, some variations from norm prototypes are perceived as errors relative to the context in which they occur and relative to the perceptual system of the listener/viewer. Pathological linguistic use is indicated when the sender-receiver's system of coordinates of ego-demarcation is sufficiently divergent from (i.e., is in great enough conflict with) the system of coordinates which constitute the system of ego-demarcation of the receiver defined by the normative values s/he assigns to the analogous selection and combination of factor-, function-, and validity-claim values in his or her own system. These values, in turn, are arrayed in a hierarchy of importance which is partially differential in each of the sender-receivers in the speech event. In writing this, we are attempting one, albeit somewhat vague, characterization of the distinction, which we all more or less clearly intuit, between normal and pathological errors. One current bone of contention is whether or not the line of demarcation between normal and pathological errors is quantitative or qualitative, or both. It is our opinion that the ultimate answer must await the collection and analysis of much more error data than has regrettably thusfar been available, within a unified analytical framework, using a much broader concept of error, a concept of error defined in relation to the entire spectrum of exemplars of the factor-, function-, and validity-claim variables mentioned above.

2. *The notion 'error' in the organon-model of language.*

In a recent study of classical malapropisms and the creation of a mental lexicon (Zwicky 1982), a monomodal/bimodal dichotomy is proposed, the former term referring to errors instantiated in single-mode perception and

production, the latter to errors instantiated when some type of production must be coordinated with some type of perception.

<i>Monomodal</i>	<i>Bimodal</i>
(1.) Speaking	(1.) Listening and Speaking
(2.) Writing	(2.) Listening and Writing
(3.) Listening	(3.) Reading and Writing
(4.) Reading	(4.) Reading and Speaking

While this dichotomy is surely a simplification — how can one read and write without reading what one has written, or how can one read and speak without listening to what one is saying, or how can one listen without concurrently speaking in inner speech, or, even monomodally, speak without listening to oneself — it is at least a suggestive start because monomodal category (3.), Listening, and bimodal categories (1.) and (2.), Listening and Speaking and Listening and Writing, refer to both a sender *and a receiver*. Furthermore, it incorporates — without comment, it is true — writing and reading as components of the theory of grammar, a silent revolution in transformational circles, but a commonplace in educationalist, psychological, and neuropsychological inquiry (cf. Avakian-Whitaker and Whitaker 1973, Beauvois and Dérouesné 1979, Coltheart 1981, 1982, 1984, Coltheart, Patterson, and Marshall 1980, Dérouesné and Beauvois 1979, Malatesha and Whitaker 1984, Marshall and Newcombe 1973, Patterson 1981, 1982, Patterson and Kay 1982, Shallice 1981, and Warrington and Shallice 1980). Then — again without comment — it rightfully places reading and writing in the center of error analysis, alongside speaking and listening, another revolution in linguistic error-analytic circles, where one searches in vain in the major anthologies of error-analytic studies (Fromkin 1973, 1980, Cutler 1982, for errors in native-language production, and Richards 1973, for errors in second-language production) for even a single paper dealing with spelling or reading errors, although such studies have been published for years in the non-linguistic (?) literature (cf., for example, Frith 1980, Goodman 1967, 1969, Weber 1968, 1970, and the many references cited there), Avakian-Whitaker and Whitaker 1973 being a notable exception. In our opinion, this in itself is a gaping error of omission, especially in view of the fact that educational, reading, and psychological journals are riddled with contributions addressed to these issues. The study of spelling and reading errors, to say nothing of the study of normal reading, writing, and spelling competence and performance, has been defined out of

the field of linguistics by a narrow conception of grammar as consisting of syntax, semantics, and phonology, in complete oblivion to an earlier, yet no less distinguished tradition (Artymovič 1932a, 1932b, Bloomfield and Barnhart 1961, Bolinger 1946, Edgerton 1941, Fries 1961, Uldall 1939, Vachek 1933, 1939, 1942, 1948, 1972, 1973, 1976, Venezky 1970, Weir and Venezky 1968, among others), with no more than a handful of exceptions (Chomsky 1970, Halle 1969, 1972, Klíma 1972, 1975, Read 1971), the gist of all of which, excepting Read 1971, is that English orthography is an optimum system of graphic representation, because it fairly consistently maps letters onto systematic phonemes, although even this claim is not without its problems (cf. Vachek 1976, and the many reviews of *The Sound Pattern of English* (Chomsky and Halle 1968)).

Zwicky (*op cit*) goes on to remark that errors may either represent the presence of an inappropriate feature or the absence of an appropriate feature, that they may be statistically aberrant, either occurring too frequently or not frequently enough, and proposes four additional distinctions crosscutting the categories monomodal and bimodal: (1.) the linguistic basis of the error; (2.) the physical relationship between the target and the error; (3.) the intentions of the producer; and, (4.) the cause of the error. He concludes with mentioning four areas of error analysis as problematic, a judgement with which we readily concur: (1.) the definition of the relevant error categories; (2.) problems with the classification of particular examples (the solution to which is provided, we think, by the solution to (1.)); (3.) problems in the recognition of examples; and (4.) devising more controlled methods of data collection. While elsewhere Zwicky (1980) briefly deals with errors which are formally correct, i.e. have no microlinguistic basis (cf. (1.) above), notably absent is any discussion of errors resulting from misfirings in any combinations of the values substitutable for any of the variables in the organon-model of language (cf. *supra*). In fact, Zwicky's theory of error, restricted as it is to speaker, hearer, code, message, and utterance, i.e. five of the factors, is *in principle* incapable of accounting for the above-mentioned combinations, because his theory of grammar is incapable of accounting for them. To mention just one example, how can a theory of grammar which consists of just a syntactic, semantic, and phonological component account for any of Austin's (1962) infelicitous uses of performatives, saying "I challenge you to a duel", when duelling has been outlawed, or saying "I do" in a would-be wedding ceremony before a ship's captain not at sea, or writing "My house" in a last will and testament when one

owns several, or uttering “I congratulate you” when it is insincere, or “I promise you” when the promise isn’t kept. This inability to account for such phenomena is not just a question of different interests or emphasis; it is rather a question of different claims about the very nature of language, about the nature of the linguistic sign (cf. *infra*), and about the nature of language in society. Rather than adhere to a set of assumptions which are so narrow as to define much of the data out of the field, the assumptions should be either revised or abandoned in favor of a new, less narrowly constrained set.

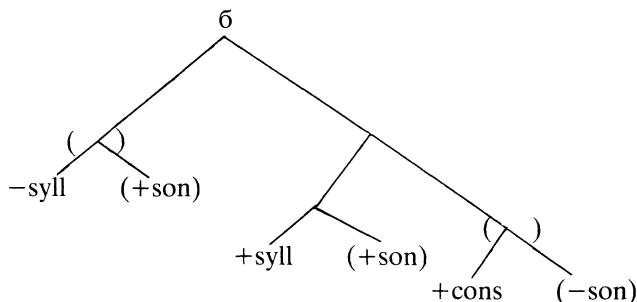
In addition to the above, one may introduce the categories of error made in native and other language acquisition — especially interesting is the negative influence of a second, rather than the first, language on a third — which Zwicky (1980, 1982) does, and errors made by bi- and multilinguals past the acquisition stage. Bankevič (1981) lucidly demonstrates the overriding influence of phonology in explaining the errors made by native speakers of Russian in the acquisition of English vocabulary. As to the question of exercising more control over the contexts of error-elicitation and observation, a significant step forward has been taken by Krashen (1982:104-112).

Most importantly, we place the study of linguistic error in the context of the entire array of linguistic factors, functions, and validity claims, in both the normal and pathological case, for if any one general conclusion has to be drawn from the studies of linguistic error conducted by linguists over the past decade, it is, and this confirms an observation made by Jakobson (1971a, 1971b, 1971c, 1971d) and Lurija (1976) years ago, that the functioning of normal language cannot *in principle* be understood without recourse to error data. How many arguments advanced by transformational linguists over the past 30 years could have been formulated without dozens of citations of the ungrammatical (*) instance, without starredness? Granted the validity of this conclusion, and one needs but become involved in an in-depth error analysis to convince oneself, the study of linguistic error, even if interest is not focused on error typology and error causation, must by definition form a weighty cornerstone of the empirical foundations of linguistics. Our attention thus shifts from an attempt to define the notion ‘possible sentence of L_i ’ to the attempt to define the notion ‘possible error of L_i in C_i ’, where C = context and is given by the organon-model of language.

3. *The notion ‘possible error’.*

Despite the veritable mushrooming of publications on the subject of error analysis during this past decade, no one, to the best of our knowledge, has addressed what we consider to be the central issue in the theory of linguistic error, namely defining the notion ‘possible error’. Luelsdorff (1986) contains such an attempt with respect to spelling errors and orthographic errors, where by ‘spelling error’ we mean an error in the product of the process of spelling and by ‘orthographic error’ we mean an error in the underlying set of orthographic conventions of which the spelling patterns in the products in part are functions (cf. Vachek 1973 for a clear review of the distinction between *orthography* and *spelling*).

Prior to introducing the concept of possible error, we dwell for a moment on the notion of ‘possible in L_i ’, heavily relying on the concept of ‘template’ lucidly discussed by Selkirk in one of her interesting works on prosodic structure. Selkirk (1980:8) points out that the issue of what a syllable looks like in a phonological representation is logically quite different from the issue of how a grammar is to give expression to the notion of ‘possible syllable in L_i ’, i.e. that individual representations of syllables do not state generalizations about syllable structure, no more than any representation of a linguistic object is a generalization about the possible forms that that linguistic object may assume. Assuming that the task of a grammar is to provide a statement of the notion ‘possible syllable of L_i ’, Selkirk supposes “that for each language this statement is in the form of a template and an accompanying set of phonological constraints... These together specify all the possible syllable types of the language and can be thought of as well-formedness conditions on the syllable structure of the phonological conditions of a language.” Equipped with this understanding of ‘template’ and phonological constraints, which together serve as well-formedness conditions, Selkirk (1980:8) exemplifies the notions by The English Syllable Template:



The function of the template is to encode the gross characteristics of syllable structure: (1.) the syllable segment-types, (2.) the order of the syllable segment-types, (3.) the hierarchical structure of the syllable segment-types, and (4.) the optionality of syllable segments and syllable-segment sequences. Selkirk summarizes by stating that the function of the template is to *contribute* to defining the well-formedness of the syllable structure of particular phonological representations, a necessary (but not sufficient) condition on which being that the syntactic structure of the representation be not distinct from the template, the sufficient condition that the collocational restrictions of the language not rule out the well-formedness of the syllable structure. Selkirk maintains that a distinction must be made between collocational restrictions and templates, because collocational restrictions do not find a “natural” expression in templates. Now, aside from the complications with which the notion ‘non-distinct’ bristles (cf. Lightner 1963, Stanley 1967) and the apparently legitimate question of why segment-structure conditions and sequence-structure conditions fail the test of generalizability in Selkirk’s sense (for the origin of the notions ‘segment-structure condition’ and ‘sequence-structure conditions’ cf. Halle 1959, where the latter notion was referred to as ‘morpheme-structure condition’, and Luelsdorff 1975, which, to the best of our knowledge, remains the only attempt to formulate such conditions for the entire consonantal inventory, including clusters, of a variety of English or any other language, for that matter), we cannot but agree that the true generalizations about a language should be embodied in an appropriate formalism, appropriate in the sense that conceptual complexity should be directly reflected in formal complexity and vice versa.

4. *The sign-template and linguistic error.*

In this section we show that linguistic errors are derivative of the formal and substantive properties of the normative templates with which they, as erroneous attempts, are correlated as targets, and the collocational restrictions placed by the given language on the elements in those templates, together with the basic transformational operations which must be posited to account for formal properties of grammars on independent grounds.

We continue with an exposition, critical examination, and revision of Mel'čuk's concept of linguistic₁ sign (relative to a language L_i, henceforth omitted), presented and explained in Mel'čuk 1982:40-41 as follows, where linguistic₁ (adj.) (<language (n.)) and linguistic₂ (adj.) (<linguistics (n.)):

A LINGUISTIC₁ SIGN: *X is a linguistic₁ sign* = X is an ordered triple =

$$X = \{Y; Z; W\},$$

where Y is the signifiant of Z, Z is the signifié of Y, and W is the syntactics of the pair {Y; Z}. Mel'čuk notes (p. 24) that the notions 'signifiant', 'signifié', and 'syntactics' are elementary, and (p. 26) that the term 'syntactics' was coined by him to denote the set of specifications (i) about all possible combinations of a given pair {signifiant, signifié} with all similar pairs and (ii) about the 'behavior' of signifiants within these combinations, these specifications being such that they cannot be deduced either from signifiants or from signifiés alone. To quote: "...the syntactics of an item describes the whole bulk of its NON-STANDARD collocability", where by 'non-standard' he means that collocability of a linguistic₁ item which cannot be stated in terms of its signifiant or its signifié, such as parts of speech, grammatical gender, verb government, conjugation or declension type, specifications of the alternations possible within the signifiant of an item, specifications which a given lexical item can condition in a neighboring item, and lexical functions (words and phrases bound by the given stem phraseologically and which stand to it in some constant semantico-syntactic relations).

We take exception to Mel'čuk's claim (p. 41) that sounds (= phones) or phonemes are examples of linguistic₁ items which are not signs, agreeing rather with Jakobson that phonemes are signifiants whose signifiés are "othernesses" so that English /p/ and English /b/, for example, each consist of a different signifiant with a different otherness as its signifié, the other-

ness of /p/ ($O_{/p/}$) and the otherness of /b/ ($O_{/b/}$), each having its own *unique* phonotactics. Since the collocability of the phonemes of English cannot be stated either in terms of their signifiants or in terms of their signifiés, however, i.e. the collocability of /p/, for example, cannot be stated in terms of either its signifiant (the classificatory distinctive feature matrix defining /p/) or its signifié (the otherness which is unique to /p/, i.e., $O_{/p/}$), we must conclude that, by definition, each of the phonemes of English has its own (unique) syntactics, in precisely Mel'čuk's use of the term 'syntactics' (cf. *supra*), which define the totality of its non-standard collocability. On this theory, phonemes, which are simple signs in the sense that they are incapable of being segmented into constituent signs, combine in accordance with their syntactics to form more complex signs called 'allomorphs'. We call signs whose signifiés are not themselves signs of other signifiés 'first-order signs'. Examples of first-order signs are the English phoneme /p/ and the English proper name /pit/ 'Pete'.

Next, we introduce the notion 'grapheme' into our theory of linguistic₁ signs, defining it as the ordered pair

$$G = \{\{L, P\} \{S_{L/P}\}\},$$

subject to revision below, where 'L' stands for the graphemic signifiant of P, 'P' the (morpho)phonemic signifié of L, and 'S_{L/P}' the syntactics of the grapheme-phoneme correspondence L/P. G consists of a signifiant, L, and a signifié, P, and a collocability which cannot be stated in terms of its signifiant, but can be stated in terms of its signifié (assuming a theory of marking conventions which mark the words in the lexicon for their origin and mark words with lexical spellings accordingly), for in the unmarked or standard case, the graphotactics of L_i is identified with the phonotactics of P_i, where L_i → P_i is a GPC of English. A linguistic₁ sign (of a language L_i) consisting of a signifié which is itself a linguistic₁ sign we call a 'second-order sign'. Since G consists of a signifié (P) which is itself a signifiant of a linguistic₁ sign whose signifié is otherness (distinctiveness), G is a second-order linguistic₁ sign.

We have maintained that in the standard or unmarked case, the graphotactics of L_i is identified with the phonotactics of P_i, where L_i → P_i is a GPC (of English). As a result, we only need to specify the phonotactics of P_i in order to obtain the graphotactics of L_i. We conclude that the normal or unmarked case of a second-order linguistic₁ sign is the case of a linguistic₁ sign with a second-order (derived) syntactics. Moreover, if in the

unmarked case the relation between grapheme and phoneme is one-to-one, we are dealing with an alphabet.

In a good many languages (e.g. English, German, Russian, French, Italian), the letters have names in the ordinary language which almost always bear an echoic relation to the (morpho)phonemes the letters are used to represent, such that the (morpho)phonemes the letters are used to represent are derivable from the names of the letters by a simple set of rules applied to the (morpho)phonemic representations of the names. For example, we derive the letter-sound of the English letter $\langle p \rangle$, i.e. [p], by first applying a rule which epenthizes the vowel *e* to all the consonantal letter-names (the names of $\langle q \rangle$, $\langle r \rangle$, $\langle w \rangle$, and $\langle y \rangle$ excepted) in the environments # ____ C and C ____ # and then applying a rule which raises the vowel *e* in the environment C ____ # to *i* (the names of $\langle k \rangle$ and $\langle j \rangle$ excepted), yielding the name [pi:] for the English letter $\langle p \rangle$ from the underlying representation /p/ via *pe*. Note that this second phonological rule does not apply in the environment # ____ C, leaving the phonetic representation of the name of the letter $\langle h \rangle$ (systematic phonemic /č/), for example, by application of just Rule I, [eč]. Two strategies for letter-sounding suggest themselves on this analysis: (1.) deriving the sounds of the letters from the surface phonetic representations of the letter names by a simple rule of vowel deletion, or (2.) suspending the application of Rules I and II, so that the morphophonemic representations of the letter names are directly identified with the phonetic representations of the letter sounds.

We have thusfar discussed the notions ‘sign’, ‘signifiant’, ‘signifié’, ‘syntactics’, ‘first-order sign’, ‘second-order sign’, ‘second-order syntactics’, ‘phoneme’, ‘grapheme’, ‘otherness’, and ‘letter name’ vs. ‘letter sound’ and proposed two alternative derivations of letter-sounds from underlying letter-names. On the second solution, the phonemic representations of the letter names, i.e. the signifiants of the letter-names, are identified with the signifiés of the letters, i.e. the signifiants of the phonemes, creating, as it were, grapheme-phoneme correspondences, although historically speaking, we might conjecture, it was the sounds of the letters which determined the selection of the letter-names on, in the vast majority of cases, the condition that the name chosen improperly or properly contain the sound the letter was being used to represent and, in the case of some orthographies (for example, Hebrew), the additional condition that the name selected be an already existing word in the language in question, this double semiotic constraint being lost in the course of time when the letters of one language

were borrowed by users of another (for example, the adoption of the Hebrew alphabet by the Greeks) (cf. Venezky 1975 for aspects of the fascinating story of letter-naming). As a consequence, the notion 'letter' or 'grapheme' is derivative in a new, perhaps more interesting sense, namely, not only is spelling a secondary system of representation in the sense that speech is phylogenetically prior to spelling, or even in the sense that speech is ontogenetically prior to writing, but in the further sense that letter-names are third-order signs, signs with graphemes as their signifiés, graphemes, in turn, having phonemes as their signifiés, and phonemes, in turn, having othernesses as their signifiés, with the interesting additional attribute of being, qua signifiants, identical to their signifiés twice removed. Indeed, letter-names and letter-sounds appear to be generic *Wahlverwandtschaften*.

Under such circumstances it is not hard to imagine why learners of English as a foreign language experience some difficulty in acquiring the names of the English letters, where the fracture is compounded in the case of learners whose native orthographies contain letters whose names are quite similar to those of the letters of English, for example, Germans. A German, under the above analysis, must learn to add Rule II which realizes the name of the letter <p> as [pi:] from *pe* (from underlying /p/) to produce the name of the letter <p> in English and to suspend the application of the rule when producing the name of the letter <p>, i.e. [pe] in German.

A simple diagram may facilitate the perception of the relationships discussed above, dealt with at greater length in Luelsdorff 1986: Chapter II:

Letter Name	/p/	3
Letter	<i>p</i>	2
Phoneme	/p/	1
Otherness	'O'	0

While a letter name is a signifiant in relation to a letter-signifié, a letter is a signifiant in relation to a phoneme-signifié, and a phoneme is a signifiant in relation to an otherness. Thus, depending upon one's perspective, a particular letter and a particular phoneme are now signifiants, now signifiés. Adopting a simultaneous view from both perspectives, which is itself *a fortiori* a perspective, we may say that letters (= graphemes) and phonems are at once signifiants/signifiés. Now, of course, we may refer to a letter name rather than use it to refer to a latter, use a letter name autonymously to refer to itself, in which case we are using a fourth-order sign, and so on, *ad*

infinitum, limited only by the upper-bound on our power to abstract.

Note that the signing relation (semiosis) between 3, 2, and 1 is not transitive. In using a letter-name to refer to a letter, where the letter refers to the phoneme (i.e. is a grapheme), we are not using the letter name to refer to the phoneme, any more than in using the name of the word ‘planet’ to refer to the word ‘planet’, where the word ‘planet’ refers to a particular planet, are we referring to a particular planet. Nor is the relation between 3 and 2 or the relation between 2 and 1 symmetric, for neither is the use of a letter-name to refer to a letter the same as using a letter to refer to a letter-name, nor the use of a letter to refer to a phoneme the same as the use of a phoneme to refer to a letter, even though the name of the letter and the sound of the letter have identical signifiants, because they do not have identical signifiés. Therefore, the relation between 3 and 2 and the relation between 2 and 1 cannot be considered equivalent. Adopting the perspective of signifiant/signifié simultaneity (cf. *supra*), moreover, the relation between {3, 1} and {1, 3}, where, qua signifiés, $1 \neq 3$, is a function, since the elements of f qua signifiés are ordered pairs with different first coordinates. From the perspective of signifiant/signifié simultaneity, the ordered triple {3, 2, 1} is identified with the ordered triple {1, 2, 3}, however, since $1 = 3$ qua signifiants, ‘top-down’ with ‘bottom-up’, as it were, the triad {name of letter, letter, (morpho)phoneme} identical to the triad {(morpho)phoneme, letter, name of letter}, which is the perspective adopted in Luebsdorff 1986, where the notation ‘GPC’ (= grapheme-phoneme correspondence) is not unreflectingly used to refer to what other writers have referred to by ‘PGC’ (= phoneme-grapheme correspondence), i.e. in total oblivion to the putative distinction between GPC’s and PGC’s, but as a convenient abbreviatory device for the perspective adopted.

At this juncture the wary reader might be concerned about the “psychological reality” of letter names in the context of *reading*, possibly granting the appropriateness of the notion in the context of spelling. The stark truth of the matter is that letter-names *are* psychologically real, in the sense that they are an inextricable part of both what might be called the ‘*faculté de lire*’ and the reading process itself, for there exist cases of letter-by-letter reading, cases of readers who cannot read without recourse to either covert or overt pronunciation of letter-names (cf. Coltheart 1984, Patterson 1981, Patterson and Kay 1982, Warrington and Shallice 1980). Surely whether or not it is possible for anyone reading English to read without at least covert recourse to letter naming is a difficult proposition to

test, but we shall assume, on the strength of the evidence cited in the works above, the essential correctness of the notion of a link between visual and echoic memory pending evidence to the contrary. Under the assumption of such a bond between letter and sound, at this level of simultaneous abstraction, the putative distinction between GPCs and PGCs is spurious, for the distinction only obtains, if it obtains at all, in real time, i.e. on-line.

Returning to the discussion of the linguistic₁ sign, we now claim that the notion ‘erroneous sign’ is derivable in all its essentials from the notion ‘sign’ itself, for plainly it is only possible for an error to obtain in relation to a normative concept whose defining elements and relations admit of such a projection. Granting the truth of this proposition, we go on to make the claim that no elements or relations in an error are novel with respect to the elements and relations contained in the normative structures themselves (concepts defined by specifying their constituents and the relations among these constituents), i.e. that the set of erroneous projections is closed under a restricted set of transformations needed in the grammar on independent grounds. Call the former proposition “The Necessary Condition on Error” and the latter “The Sufficient Condition on Error”. The former proposition expresses the notion that the form of the actual is preshaped, so to speak, by the form of the possible, for clearly something cannot exist which is in principle excluded from the realm of existence; the latter proposition expresses the notion that the elements and relations of the norm, on the one hand, and the templates and rules which specify them, on the other, predetermine all the elements and relations of the norm-deviations, to the extent that, in the unmarked case, the norm-deviations contain no elements or relations which are not present in the templates underlying the form of the norm. This is not to say that one does not find new productions in the deviations from the norm — clearly, by definition, one does — but that the elements in the types of deviations qua products and the processes underlying their production are themselves norm-immanent.

5. *Sign mutation.*

We systematically mutate the elements and relations given by Mel'čuk's definition of ‘linguistic₁ sign’, which we have reconceptualized as a sign-template, assigning the template an empirical interpretation in the domain of English GPCs in order to ascertain whether the resultant mutants bear any resemblance to the spelling errors in our corpus (reproduced

in Luelsdorff 1986) and in demonstration of the validity of The Necessary and Sufficient Conditions on Error, although the conditions are intended to hold for all empirical domains of error in all languages in the unmarked or core case (cf. Vachek 1966 for extensive discussion of the notion 'core'). We remind the reader of the discussion in section 1, where it was concluded that the error categories one adopts are a function of the system of coordinates from which the errors are viewed.

Table I
Error Sign-Template

<i>Error Type</i>	$X = \{Y, Z, W\}$	<i>Denomination</i>
	+ + +	Norm
I	+ + -	x
II	+ - -	x
III	- - -	x
IV	- + +	x
V	- + -	x
VI	+ - +	x
VII	- - +	x

where + = normative representation and - = non-normative representation. It should be apparent that Mel'čuk's notion of linguistic₁ sign is inadequate to the task of serving as an error sign-template, although it was far from obvious that it was inadequate as a sign-template upon consideration of his detailed discussion and exemplification of at least 50 different types of normative linguistic₁ signs. The heart of the problem is that his theory of the linguistic₁ sign tolerates neither spurious *absence* of a signifiant and/or signifié, as in errors of omission, nor the spurious *presence* of a signifiant and/or signifié, as in errors of addition, nor spurious paradigmatic representation, as in errors of substitution, nor spurious syntagmatic representation, as in errors of displacement and transposition. In short, it does not tolerate the dissociation of a signifiant from a signifié, because his notion of syntactics is restricted to the ordered pair {Y, Z}. Consequently, we propose the following revision of Mel'čuk's definition of 'linguistic₁ sign' in language L_i:

$$X = \{\{Y, S_Y\} \{Z, S_Z\}\},$$

which is to be read: a linguistic₁ sign X is identified with the ordered pair consisting of the ordered pair {signifiant Y of X, syntactics S of Y}, on the

one hand, and the ordered pair {signifié Z of X, syntactics S of Z}, on the other. Here we are no longer assigning syntactics to the signifiant/signifié qua *pair*, but to the signifiant and to the signifié qua *members* of the pair. Furthermore, in order to distinguish between errors of addition, which involve the erroneous presence of an item, and errors of omission, which involve the erroneous absence of an item, we supplement our coefficients + and - with the coefficient Ø, where + = normative representation, - = non-normative representation, and Ø = absence of representation. Since the theory may now distribute the operation of dissociation over the operation of semiosis, and dissociation may now entail the erroneous presence or the erroneous absence of signifiers and/or signifieds, we call this revision "The Distributive Dissociative Error Sign-Template", represented here, together with a partial array of the errors it predicts, in Table II:

Table II
Distributive Dissociative Error Sign-Template

<i>Error Type</i>	X = ((Y, S _Y) (Z, S _Z))	<i>Denomination</i>
	+	Norm
I	- - + +	LA
II	Ø Ø + +	LO
III	- + + +	LS
IV	+ - + +	LD
V	+ + - -	PA
VI	+ + Ø Ø	PO
VII	+ + - +	PS
VIII	+ + + -	PD
IX	- - - -	LA + PA
X	Ø Ø Ø Ø	LO + PO
XI	- + - +	LS + PS
XII	+ - + -	LD + PD
XIII	- - Ø Ø	LA + PO
XIV	- - - +	LA + PS
XV	- - + -	LA + PD

where L = Literal, P = Phonological/Phonetic, A = Addition, O = Omission, S = Substitution, and D = Displacement.

6. Summary.

The study of linguistic error has been traditionally limited to the investigation of non-normative features of the linguistic code, message, and utterance. We suggest an extended view of linguistic error, broad enough to encompass all the factors, functions, and validity claims of speech events within the organon model of language.

Central to the theory of linguistic error is the notion ‘possible linguistic error’. Since a language is a system of signs, an inquiry into erroneous language is an inquiry into erroneous signs. The Necessary Condition on Error states that the notion ‘erroneous sign’ is derivable in all its essentials from the notion ‘sign’ itself, since deviations necessarily only obtain in relation to the regularities which underlie them. The Sufficient Condition on Error holds that the preexisting elements and relations in the norm are sufficient to specify the elements and relations in the deviations (cf. Buszkowski & Luelsdorff 1986 for extensive formalization).

The traditional notions of a linguistic sign as the relation between a signifiant and a signifié or as the relation between a signifiant and a signifié and its associated syntax are seen to be inadequate to the task of serving as an error sign-template. In their stead, we urge a dissociative view of the linguistic sign, one permitting the dissociation of a signifiant from a signifié, or a signifié from a signifiant, their spurious substitution, addition, and omission.

BIBLIOGRAPHY

- Artymovič, Agenor. 1932a. “Fremdwort und Schrift.” In: *Charisteria Gu. Mathesio quinquagenario...oblata*, Prague.
- Artymovič, Agenor. 1932b. “Pysana mova,” *Naukovyj zbirnik ukrainskoho Vys. Ped. Instituta v Prazi II*, 1-8.
- Austin, John. 1962. *How to do things with words*. Cambridge, Mass.: Harvard University Press.
- Avakian-Whitaker, H. and Whitaker, H.A. 1973. “The spelling errors of children with communication disorders: A preliminary classification,” *Linguistics*, 115, 106-118.
- Bankevič, Leonid Valentinovič. 1981. *Testirovanie leksiki inostrannogo jazyka* [Testing a foreign language lexicon]. Moscow: Vysšaja škola, pp. 112.

- Beauvois, Marie-France and J. Dérouesné. 1979. "Phonological dyslexia: Three dissociations," *Journal of neurology, neurosurgery, and psychiatry* 42, 1115-1124.
- Bloomfield, Leonard and Clarence Barnhart. 1961. *Let's read: A linguistic approach*. Detroit: Wayne State University Press.
- Bolinger, Dwight. 1946. "Visual morphemes," *Language* 22, 333-340.
- Buszkowski, Wojciech & Philip A. Luelsdorff 1986. "A formal approach to error taxonomy." In: *Language and discourse: Test and protest. A Festschrift for Petr Sgall*, edited by Jacob L. Mey. Linguistic & Literary Studies in Eastern Europe. Vol. 19. Amsterdam: John Benjamins B.V., pp. 217-242.
- Bühler, Karl. 1920. "Kritische Musterung der neueren Theorien des Satzes," *Indogermanisches Jahrbuch*, Vol. VI, Berlin/Leipzig, 1-20.
- Bühler, Karl. 1931. "Phonetik und Phonologie." In: *Travaux du cercle linguistique de Prague* 4. *Réunion phonologique internationale tenue à Prague* (18-21/XII, 1930). Prague: Jednota československych matematiku a fysiká, 22-53.
- Bühler, Karl. 1934. *Sprachtheorie: Die Darstellungsfunktion der Sprache*. Jena: Gustav Fischer.
- Bühler, Karl. 1936. "Das Strukturmodell der Sprache." In: *Travaux du cercle linguistique de Prague* 6. *Etudes dédiées au quatrième congrès des linguistes*. Prague, 3-12.
- Bühler, Karl. 1969. *Die Axiomatik der Sprachwissenschaften*. Introduction and commentary by Elizabeth Ströker. Frankfurt/Main: Vittorio Klossermann.
- Chomksy, Carol. 1970. "Reading, writing, and phonology," *Harvard educational review* 40, 289-309.
- Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
- Coltheart, Max. 1981. "Disorders of reading and their implications for normal reading," *Visible language* 15:3, 245-286.
- Coltheart, Max. 1982. "The psycholinguistic analysis of acquired dyslexias: Some illustrations," *Philosophical transactions of The Royal Society*, B. London.
- Coltheart, Max. 1984. "Acquired dyslexias and normal reading." In: R.N. Malatesha and H.A. Whitaker (eds.) *Dyslexia: A global issue*. The Hague: Martinus Nijhoff, 357-374.
- Coltheart, Max, Karalyn Patterson, and John Marshall (Eds.) 1980. *Deep dyslexia*. London: Routledge & Kegan Paul.

- Cutler, Anne. 1982. *Speech errors: A classified bibliography*. Distributed by the Indiana University Linguistics Club, pp. 34.
- Dérouesné, J. and Marie-France Beauvois. 1979. "Phonological processing in reading: Data from alexia," *Journal of neurology, neurosurgery and psychiatry*, 42, 1125-1132.
- Edgerton, W.F. 1941. "Ideograms in English writing," *Language* 17, 148-150.
- Fries, Charles C. 1961. *Linguistics and reading*. New York: Holt, Rinehart and Winston.
- Frith, Uta (Ed.) 1980. *Cognitive processes in spelling*. London: Academic Press.
- Fromkin, Victoria A. (Ed.) 1973. *Speech errors as linguistic evidence*. The Hague: Mouton.
- Fromkin, Victoria A. (Ed.) 1980. *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*. New York: Academic Press.
- Garrett, Merrill F. 1981. "The organization of processing structure for language production: Applications to aphasic speech." Revised version of a paper presented at the Conference on Biological Perspectives on Language, Montreal, May 17-20, 1981.
- Goodman, Kenneth S. 1967. "Reading: A psycholinguistic guessing game," paper read at the American Educational Research Association, New York.
- Goodman, Kenneth S. 1969. "Analysis of oral reading miscues: applied psycholinguistics," *Reading research quarterly* 5:1, 9-30.
- Habermas, Jürgen. 1973. *Legitimationsprobleme im Spätkapitalismus*. Frankfurt/Main: Suhrkamp.
- Habermas, Jürgen. 1976a. *Technik und Wissenschaft als "Ideologie"*. Frankfurt/Main: Suhrkamp.
- Habermas, Jürgen. 1976b. *Zur Rekonstruktion des Historischen Materialismus*. Frankfurt/Main: Suhrkamp.
- Habermas, Jürgen. 1976c. "Universalpragmatische Hinweise auf das System der Ich-Abgrenzung." In: Auwärter, Manfred et al. (Eds.) *Seminar. Kommunikation. Identität*. Frankfurt/Main: Suhrkamp, 332-347.
- Habermas, Jürgen. 1976d. "Was heißt Universalpragmatik?" In: Apel, Karl-Otto (ed.) *Sprachpragmatik und Philosophie*. Frankfurt/Main: Suhrkamp, 174-272.
- Halle, Morris. 1959. *The sound pattern of Russian*. The Hague: Mouton.
- Halle, Morris. 1969. "Some thoughts on spelling." In: K.S. Goodman and J.T. Fleming (Eds.) *Psycholinguistics and the teaching of reading*.

- Newark, Delaware: International Reading Association.
- Halle, Morris. 1972. "On a parallel between conventions of versification and orthogrpahy; and on literacy among the Cherokee." In: J.F. Kavanagh and I.G. Mattingly (Eds.) *Language by eye and ear*. Cambridge, Massachusetts: The MIT Press, 149-154.
- Holenstein, Elmar and Tarcisius Schelbert (Eds.) 1979. *Roman Jakobson Poetik: Ausgewählte Aufsätze 1921-1971*. Frankfurt/Main: Suhrkamp.
- Hymes, Dell H. 1961. "Functions of speech: An evolutionary approach." In: H. Gruber (Ed.) *Anthropology and education*. Philadelphia: University of Pennsylvania Press.
- Hymes, Dell H. 1968. "The ethnography of speaking." In: Gladwin and Sturtevant (Eds.) *Anthropology and human behavior*. Washington, D.C.: The Anthropological Society of Washington. Reprinted in: Fishman (Ed.) *Readings in the sociology of language*, 99-138.
- Jakobson, Roman O. 1960. "Closing statement: Linguistics and poetics." In: Thomas A. Sebeok (Ed.) *Style in language*. Cambridge, Massachusetts: MIT Press, 350-377.
- Jakobson, Roman O. 1971a. "Zwei Seiten der Sprache und zwei Typen aphatischer Störungen." Reprinted in: *Selected writings: Vol. II. Word and language*. The Hague: Mouton, 117-141.
- Jakobson, Roman O. 1971b. "Aphasia as a linguistic topic." Reprinted in: *Selected writings: Vol. II. Word and language*. The Hague: Mouton, 229-238.
- Jakobson, Roman O. 1971c. "Toward a linguistic classification of aphasic impairments." Reprinted in: *Selected writings. Vol. II. Word and language*. The Hague: Mouton, 289-306.
- Jakobson, Roman O. 1971d. "Linguistic types of aphasia." Reprinted in: *Selected writings. Vol. II. Word and language*. The Hague: Mouton, 307-333.
- Klima, Edward S. 1972. "How alphabets might reflect language." In: J.F. Kavanagh and I.G. Mattingly (Eds.) *Language by eye and ear: The relationships between speech and reading*. Cambridge, Massachusetts: The MIT Press, 57-80.
- Klima, Edward S. 1975. "Sound and its absence in the linguistic symbol." In: J.F. Kavanagh and J.E. Cutting (Eds.) *The role of speech in language*. Cambridge, Massachusetts: The MIT Press.
- Krashen, Stephen D. 1982. *Principles and practice in second language acquisition*. New York: Pergamon Press.
- Lightner, Theodore M. 1963. "A note on the formation of phonological

- rules," *Quarterly progress report*, Research Laboratory of Electronics, MIT, No. 68, 187-189.
- Luelsdorff, Philip A. 1975. *A segmental phonology of Black English*. Janua Linguarum, Series Practica, 191. The Hague: Mouton.
- Luelsdorff, Philip A. 1982. "Differential linguistics: A program for research," *Folia phoniatrica* 34: 173-181.
- Luelsdorff, Philip A. 1984. "Einstellung als ein Faktor des erweiterten Organon-Modells der Sprache." In: Achim Eschbach (Ed.) *Bühler-Studien*. Vol. I. Frankfurt/Main: Suhrkamp, 317-342.
- Luelsdorff, Philip A. 1986. *Constraints on error variables in grammar: Bilingual misspelling orthographies*. Amsterdam: John Benjamins.
- Luria (= Lurija), Alexandr Romanovič. 1976. *Basic problems of neurolinguistics*. The Hague: Mouton.
- Malatesha, R.N. and Harry A. Whitaker (Eds.) 1984. *Dyslexia: A global issue*. NATO ASI Series, Series D, Behavioural and Social Sciences: No. 18. The Hague: Martinus Nijhoff.
- Marshall, John C. and Freda Newcombe. 1973. "Patterns of paralexia," *Journal of psycholinguistic research* 2, 175-199.
- Maxwell, Edith M. 1982. *A study of misarticulation from a linguistic perspective*. Indiana University Linguistics Club.
- Mel'čuk, Igor' A. 1982. *Towards a language of linguistics. A system of formal notions for theoretical morphology*. Internationale Bibliothek für allgemeine Linguistik/International library of general linguistics, Eugenio Coseriu (Ed.), Vol. 44, revised and edited by Philip A. Luelsdorff, Munich: Wilhelm Fink Verlag.
- Patterson, Karalyn E. 1981. "Neuropsychological approaches to the study of reading," *British journal of psychology*, 72, 151-174.
- Patterson, Karalyn E. 1982. "The relation between reading and phonological coding: Further neuropsychological observations." In: Andrew W. Ellis (Ed.) *Normality and pathology in cognitive functions*. London: Academic Press, 77-112.
- Patterson, Karalyn E. and Janice Kay. 1982. "Letter-by-letter reading: Psychological descriptions of a neurological syndrome," *Quarterly journal of experimental psychology* 34A, 411-441.
- Read, Charles. 1971. "Preschool children's knowledge of English phonology," *Harvard educational review*, 41, 1-34.
- Richards, Jack C. (Ed.) 1973. *Error analysis: Perspectives on second language acquisition*. London: Longman.
- Selkirk, Elizabeth O. 1980. "On prosodic structure and its relation to syn-

- tactic structure," Indiana University Linguistics Club.
- Shallice, T. 1981. "Neurological impairment of cognitive processes," *British medical bulletin*, 37, 187-192.
- Stanley, Richard. 1967. "Redundancy rules in phonology," *Language*, 43, 2, 393-436.
- Stemberger, Joseph P. 1983. "Speech errors and theoretical phonology: A review," Indiana University Linguistics Club.
- Uldall, H.J. 1939. "Speech and writing," *Congrès international des sciences anthropologiques et ethnologiques, Compte rendu de la deuxième session*. Copenhagen.
- Vachek, Josef. 1933. "Český pravopis a struktura čestiny," *Listy filologická* 60.
- Vachek, Josef. 1939. "Zum Problem der geschriebenen Sprache," *Travaux du cercle linguistique de Prague* 8, 94-104.
- Vachek, Josef. 1942. "Přesný jazyk a přavopis," *Čtení o jazyce a poesii* 1.
- Vachek, Josef. 1948. "Written language and printed language," *Recueil linguistique de Bratislava* 1, 67-74.
- Vachek, Josef. (Ed.) 1966. *Les problèmes du centre et de la périphérie du système de la langue. Travaux linguistique de Prague* 2. Prague: Academia.
- Vachek, Josef. 1972. "The present state of research in written language," *Folia linguistica* 6, 47-61.
- Vachek, Josef. 1973. *Written language*. Janua Linguarum, Series Critica, 14. The Hague: Mouton.
- Vachek, Josef. 1976. *Selected writings in English and general linguistics*. The Hague: Mouton.
- Venezky, Richard L. 1970. *The structure of English orthography*. Janua Linguarum, Series Minor, 82. The Hague: Mouton.
- Venezky, Richard L. 1975. "The curious role of letter names in reading instruction," *Visible language* 9:1, 7-23.
- Warrington, E. and T. Shallice. 1980. "Word-form dyslexia," *Brain* 103, 99-112.
- Weber, Rose-Marie. 1968. "The study of oral reading errors: A survey of the literature," *Reading research quarterly* 4:1, 96-119.
- Weber, Rose-Marie. 1970. "A linguistic analysis of first-grade reading errors," *Reading research quarterly* 5:3, 427-451.
- Weir, Ruth and Richard L. Venezky. 1968. "Spelling-to-sound patterns." In: K.S. Goodman (Ed.) *The psychological nature of the reading process*. Detroit: Wayne State University Press.

- Zwický, Arnold M. 1980. *Mistakes*. Reynoldsburg, Ohio: Advocate.
- Zwický, Arnold M. 1982. "Classical malapropisms and the creation of the mental lexicon." In: Loraine K. Obler and Lise Menn (Eds.) *Exceptional language and linguistics*. New York: Academic Press.

A PRELIMINARY REPORT ON A PILOT INVESTIGATION OF GREENLANDIC SCHOOL CHILDREN'S SPELLING ERRORS

Birgitte Jacobsen

1. *Introduction.*

In 1973 there was a rather radical spelling reform in Greenland from an etymological spelling to a phonemic one, based on the Central West Greenlandic form of the language. To my knowledge the only other investigation of Greenlandic spelling at all is an unpublished report (Petersen, 1970) of an experiment on different ways of indicating length. This served as a preparation for the spelling reform.

This investigation is the first one after the spelling reform. It is a pilot project the purpose of which is to map out which spelling errors school children actually make in the new phonemic orthography. Areas of specific interest which can serve as starting points for further investigations will then be isolated. Some cognitive processes are suggested, but as this is a pilot investigation I will not consider problems of bilingualism or the acquisition of spelling a polysynthetic language here. As far as information has been available I have weeded out children speaking other dialects, as well as children with severe spelling and reading troubles, in order to have a homogenous data sample. Later I found that this latter step might not have been necessary (Frith 1980, Moats 1983, Nelson 1980). The data was extracted from several hundred dictations, collected from two schools in Nuuk, the main town. It includes dictations exclusively, because my mastery of the Greenlandic language at the time of collection was insufficient for me to identify spelling errors in essays. Nine different classes are represented, from grade 3 to grade 8. The total number of errors is 5,600.

2. *The phoneme inventory and the two spellings.*

Greenlandic is a polysynthetic language, and in the elaboration of complex words major morphological and phonetic changes take place, at morpheme boundaries as well as morpheme-internally. Far-ranging processes of assimilation have levelled almost all consonant clusters and diphthongs. So the old etymological orthography's preservation of information as to underlying forms has no reality for the present-day speaker. Greenlandic has a length distinction in vowels and consonants (except for /j/). The three vowels have wide-ranging phonetic manifestations, the most marked allophony being pharyngealization before uvular consonants (/q/, /r/).

The phoneme inventory:

a	i	u	
<hr/>			
p	t	k	q
m	n	ŋ	
v	l	g	r
		j	
		s	

/v/ = [v, w] /g/ = [ɣ] /r/ = [ʁ]

Long voiced continuants are unvoiced: /v:/ = [f:], /g:/ = [x:], /r:/ = [χ:], /l:/ = [ɬ:]. /ŋ/ is written <ng>.

Long phonemes are represented as geminates.

The old orthography dates back to 1851, when Samuel Kleinschmidt made the first scientifically adequate description of the language. In the old spelling long vowels were represented by different accents or as diphthongs, while long consonants were represented by different accents over the preceding vowel, or as clusters, depending on the underlying forms. An illustrative example: When a child wrote about his /aallaat/ 'gun', he had to decide if the first long /aa/ should be <â>, <ai>, or <au>. When he had decided that, the problem arose whether long (voiceless) /ll/ was to be written <vdll>, <tdll>, or <gdl>. And then again he had three possibilities for the last long /aa/. 27 possibilities in all! The correct form is <autdlaɪt>. In the new orthography long phonemes, including the voiceless continuants, are written as geminates, except for long /vv/, where the orthography over-differentiates and requires <ff>. /ŋŋ/ is written <nng>. (For a discussion of whether long phonemes should be described as long or as geminates see

Rischel 1974:77-100. In accordance with his solution I denote the long phonemes by geminates.) The old spelling's overdifferentiation of the vowels, with <e> and <o> being used for the pharyngealized allophones before <r> and <q>, is preserved in the new spelling. Both spelling systems have only one <a>-grapheme, though also here the pharyngealized allophone is clearly distinct from the "plain" [a]. The old spelling also wrote <e> and <o> word-finally because of a slight lowering in this position. In the new spelling this inconsistency is removed.

There are three exceptions to the use of geminates for long segments:

(1) Word-final /ai/: [ɑ_k] is preserved as a diphthong, and appears as <ai> in the spelling.

(2) /ts/. This cluster has undergone only partial assimilation to [t^s], and this is the only instance of progressive assimilation. (For measurements of the relative durations of the occlusive phase and the sibilant phase see Mase and Rischel 1971). Long /tt/ before /i/ merges with /ts/ into [t^s] due to an automatic assibilation of /t/ before /i/. This assibilation also takes place before the pharyngealized allophone of /i/: [ɛ-] (broad transcription), though this actually is lower than the highest allophones of /a/ which cause no assibilation. In the new spelling long /tt/ before /i/ is written <ts>. This is phonemically inconsistent (Rischel 1972:6-7), as it looks odd to indicate the automatic assibilation before /i/ only in the long /tt/:

<uatsinni> /..tti../ [..tt^si..] 'at our house'

but not in the short /t/:

<qitornait> /..ti../ [..t^si..] 'your children'

Another inconsistency, however, would arise if <tti> were written instead. In certain roots /s/ is "strengthened" to /ts/ under suffixations:

<nasaq> 'cap' <natsat> 'caps'.

With the <tti> solution the "strong" /s/, /ts/, would be written differently before /i/, as for example in <tattit> 'lakes' from <taseq> 'lake'. This would be morphologically inconsistent. So, <tsi> was chosen in favour of <tti>, because the morphological inconsistency was considered more serious than the phonetic.

(3) The last consonant cluster which is not totally assimilated is the /rC/-cluster. (Rischel 1974:47: "...that there is sufficient evidence for regarding the "r"-clusters of West Greenlandic as being assimilated into long or

geminates with concomitant pharyngealization.") These /r/-clusters have a clear pharyngeal starting phase which is made more distinct due to the pharyngealization of the preceding vowel: [V-^bCC] (For convenience "pharyngealized" will henceforth be labelled "uvularized".) The /r/-clusters are rendered with explicit <r> as well as with a uvularized vowel, e.g.

<erneq> /irniq/ [e-^bnne-q] 'son'

The spelling system is thus redundant as regards the uvularity of /rC/- clusters: <erC..>, but non-redundant as regards the quantity of the second consonant of the cluster, its length being predictable: <rC> = [^bCC]. This last is due to the principle of graphic economy, and is also in accordance with the general constraint of having no clusters with three consonants (except for <nng>, /ŋŋ/). However, all (im)possibilities occur as spelling errors:

For <iserpoq>, /isirpuq/, [ise-^bppɔq], 'he went in':

- (1) *<iserppoq>
- (2) *<iseppoq>
- (3) *<isepoq>
- (4) *<isirpoq>
- (5) *<isirppoq>
- (6) *<isippoq>
- (7) *<isipoq>

As can be seen, both doubly redundant forms like (1) and non-redundant forms like (3) and (4) occur.

The last problematic issue to mention is the approximants /j/ and /v/. The automatic intervocalic glides, [j] after /i/ and [w] after /u/ are not written in the new orthography, as opposed to the old one. E.g.:

<siumut> [si^jumut] 'forward'

<puisi> [pu^wisi] 'seal'

and neither is the morphological /v/ in these environments, which causes some morphological inconsistency:

/neri/ + /vaa/ ('eat + he-it') : <nerivaa>

but /taku/ + /vaa/ ('see + he-it') : <takuaa>

Hence — since spelling of quantity always yields problems — there are five specific areas where errors are expected:

- (1) quantity
- (2) uvularized/non-uvularized vowels
- (3) <tsi>/<ti> errors
- (4) /rC/-clusters
- (5) glides/morphological approximants

In a first investigation it is necessary to consider errors in these five groups as conventional errors, the question of slips generating such errors being of less relevance. Since there are only three vowel phonemes, many vowel substitutions might alternatively be "explained" as being slips, with overt trigger. This is true, e.g. of area (2) above:

*<niqi> for <neqi> 'meat'

But then one important generalization will be missing, namely that this type of substitution of *<i> for <e> and *<u> for <o> is unique, being due to the overdifferentiating system in uvularized environments, and is thus completely different from substitutions like

*<utuartut> for <atuartut> 'school children'

The same is true as regards quantity errors. Every time a geminate is reduced, or a single segment is doubled, one could view it as a case of length assimilation (or dissimilation) to some suitable environment. But this is not just a matter of "omissions" and "additions", but a general error complex inherent in languages with a quantity distinction. (cf. e.g. Nauclér 1980 for quantity errors in Swedish.) In the following paragraphs the five groups of errors will be exemplified and discussed.

As regards my error taxonomy it is primarily a tool. Of course it took some time to make it suitable. It is segmental. Errors can unambiguously be placed within it without precluding later generalizations.

3. *The error taxonomy.*

(The examples are illustrative, not exhaustive). *<...> = error word
<...> = target word

(Type 1) Reduction of a geminate consonant.

*<iluanaarsinaasarput> (<iluanaarsinnaasarput>) 'they used to do well'

(Type 2) *Gemination of a single consonant.*

*<akkileriikkuutaarlutillu> (<akileriikkuutaarlutillu>) ‘paying by installments’

(Type 3) *Insertion of a consonant.*

- a) Glides.
*<sijuani> (<siuani>) ‘in the stem of a boat’
- b) Insertion of *<n> or *<g> to form the digraph <ng>.
*<atuangannguatit> (<atuagannguatit>) ‘the small books of yours’

This error type is characteristic of dialects with phonological substitution of /ŋ/ for /g/, but is also relevant for the general problem of digraphs. Specific tests are required for more information.

- c) Others (categorized as obvious slips).
*<anaanan> (<anaana>) ‘mother’

(Type 4) *Omission of a consonant.*

- a) Glides.
*<aperauk> (<aperajuk>) ‘did you ask him’ (NB. also *<a> for <aa>.)
- b) Omission of <n> or <g> in the digraph <ng>.
*<agut> (<angut>) ‘man’ See the remark in 3b.
- c) Others (categorized as obvious slips).
*<akera> (<akeraq>) ‘enemy’

(Type 5) *Substitution of consonants.*

- a) Glides/approximants.
*<nugakippoq> (<nujakippoq>) ‘he is thin-haired’
- b) Systematic. (e.g. *<k> for <q>, *<g> for <k> etc. cf. §6)
*<kimmeq> (<qimmeq>) ‘dog’
- c) Others (categorized as obvious slips).
*<kika> (<kina>) ‘who?’

(Type 6) *Reversal of consonants.*

*<asta> (<atsa>) ‘father’s sister’

(Type 7) <r>-insertion.

*<naterq> (<nateq>) ‘floor’

(Type 8) *<r>-omission.*

*<isepoq> (<iseroq>) ‘he went in’

(Type 9) *<s>-insertions before /i/.*

*<qitsigissat> (<qitigissat>) ‘dance music’

(Type 10) *<s>-omission in /ts/-cluster.*

*<masatinngagit> (<masatsinnagait>) ‘without making them wet’

(Type 11) *Reduction of a geminate vowel.*

*<inusaq> (<inuusaq>) ‘doll’

(Type 12) *Gemination of a single vowel.*

*<tuusaavara> (<tusaavara>) ‘I heard it’

(Type 13) *Insertion of a vowel.*

a) In the diphthong <ai>.

*<ilivaaai> (<ilivai>) ‘he placed them’

b) Others (categorized as obvious slips).

*<natiki> (<nateq>) ‘floor’

(Type 14) *Omission of a vowel.*

*<qtsuk> (<qitsuk>) ‘cat’

(Type 15) *Substitution of vowels.¹*

a) Uvularized/non-uvularized.

*<atuartot> (<atuartut>) ‘school children’

*<manniqartarfiit> (<manneqartarfiit>) ‘their egg store’

b) Others (categorized as obvious slips).

*<utuartut> (<atuartut>) ‘school children’

(Type 16) *Reversal of vowels.*

*<sailuk> (<sialuk>) ‘rain’

(Type 17) *Reversal of consonant and vowel.*

*<iupt> (<input>) ‘oar’

*<liumi> (<illumi>) ‘in the house’ (NB. The unpermitted sequence of two consonants initially is avoided, *<lli..>)

(Type 18-21) *Addition, omission, substitution and reversal of syllables.*

(Minimum entity: CV)

(Type 22-25) *Ditto for entities larger than the syllable.*

(Type 26) *Word-division errors.*

(Type 27) *Substitution of a vowel for a consonant.* (But cf. §7)

*<nerioaa> (<nerivaa>) 'he ate it'

(Type 28) *Substitution of a consonant for a vowel.* (But cf. §7)

*<ilivaj> (<ilivai>) 'he placed them'

(Type 29) *Axis orientation error.*

*<p> for <q>, e.a.

(Type 30) *Interference from the old spelling.*

Use of accents, clusters or diphthongs.

4. *Interference from the old spelling.*

The old spelling is still used, primarily by the older and the middle-aged generation. It still occurs in newspapers and books. So every child knows the old spelling passively, to a greater or lesser extent. Still I have found very few interference errors, most of them in proper names or proper name-like words, e.g.:

*<atâta> (<ataata>) 'father' In old spelling: <atâta>.

*<Nûk> (<Nuuk>) 'the cape' (Name of the main town) In old spelling: <Nûk>. Notice the funny mixing of systems.

*<nauja> (<naaja>) 'sea gull' (Also a girl's name). In old spelling: <nauja>.

But such errors also occur in other kinds of words, and sometimes with erroneous use of the old conventions:

*<iglit> (<illit>) 'you'. Wrong for old spelling's <ivdlit>.

*<nikuivpoq> (<nikuippoq>) 'he rose'. Wrong for old spelling's <nikuípok>.

The occurrence of *<e> and *<o> word-finally cannot automatically be taken as interference from the old spelling (cf. §2), since confusion of uvularized/non-uvularized vowels is a general error. Some of the final *<e>'s and *<o>'s probably are interference errors, but it will be rather difficult to decide amongst these possibilities. (But cf. §8).

5. *Interference from Danish.*

Since the Danish colonization in the 1720's a lot of Danish loanwords have come into the language. Until recently they were adapted to Greenlandic phonotactics, but since the 1950's the cultural and linguistic influence has been strongly intensified. This has caused a decreasing degree of phonetic adaption of loanwords, so now they can violate the native phonotactics severely, which manifests itself in the spelling. Furthermore the alphabet contains alien letters. All children have Danish lessons in school. In spite of this intense language contact only very few interference errors occur. In my data there occur only two substitutions of a Danish vowel, namely /å/: [ɔ] for the Greenlandic uvularized /u/: [ɔ]:

- *<anåre> (<anori>) 'wind' (plus final vowel substitution)
- *<qanåd> (<qanoq>) 'how, what' (plus axis orientation error, *<d> for <q>.)

As the new spelling is not consistent in writing <p> or , <t> or <d>, <k> or <g> in loanwords I do not count this kind of substitution as an error, when it occurs in loanwords. It also occurs, though not very frequently, in native words, especially amongst poor spellers. Here it must be taken as an interference error, either from Danish directly, or from the alphabet's containing alien letters. Danish /b/, /d/, and /g/ are unvoiced stops, as are Greenlandic /p/, /t/, and /k/, so substitution is not unmotivated. There are more occurrences of *<g> for <k> than of the other two substitution pairs. This is of course because <g> is a native letter, and some of these substitutions are hence language-internal, general substitutions. (cf. §6).

One possible type of interference of more interest is not segmental but prosodic. A specific use of geminated consonants may be taken as interference from the Danish convention of vowel-length assignment. For more details see §10.

6. *Systematic substitutions of consonants.*

It is not uncommon to substitute *<k> for <q>, even in the older school classes. This might show interference from the old spelling of /q/, which was small capital κ, <κ>. The capital was <Κ̄>.

It might also be a substitution of one place of articulation for another

(Nauclér 1980:138). Substitution of *<q> for <k> is very rare. It is not uncommon to substitute *<g> for <k> (while the opposite is rare.) This might be interference from Danish (cf. §5). It might also be a substitution with “preservation of salient feature” (Moats 1983:125). Substitutions of approximants will be considered in §7. Substitutions of *, *<d>, and *<g> for <p>, <t>, and <k> respectively, were considered in §5.

Let me present some curiosities from grade 4 and grade 6:

- *<ighu> (<iggu>) ‘how sweet you are! (<gg> = [xx])
- *<tigupaslapaa> (<tigupallappaa>) ‘he hurried to take them’ (<ll> = [ɿ])
- *<pisaqarsimasuni> (<pisaqarsimalluni>) ‘having caught game’
(NB. <l> is omitted in the substitution of *<sl> for <ll>.)
- *<kerti> (<qiti>) ‘the middle’ (plus vowel substitution *<e> for <i>.)
- *<kermeq> (<qimmeq>) ‘dog’.”

These errors probably are cases of *segmentalizations*. (Michael Fortescue has drawn my attention to the fact that *<sl> for unvoiced /l/ is a frequently occurring spelling error in some Western Canadian texts.) Of course this would be more plausible if the suggested *<kr> for initial <q> turned up somewhere without metathesis.

Another peculiar substitution is that of three children speaking the /g/-nasalization dialect, in which /ŋ/ is phonologically substituted for /g/: [ɣ]. They make errors of the following type:

- *<sinngulik> (<siggulik>) ‘provided with a bill’
- *<tinngak> (<tiggak>) ‘male seal’

Long /gg/: [xx] is *never* subject to nasalization in any dialect. So what we have here is a dialectal confusion taking place on the *graphemic surface* level. Notice the correct quantity!

7. Glides and approximants.

Because of the often very weak pronunciation of the approximants /j/, /v/, and /g/, it can be hard to hear which one it is. There is a morphological difference between them, but as mentioned in §2 the orthography also takes phonetic considerations into account when assigning them. Thus <j> is never written intervocally after /i/, and <v> is never written intervocally after /u/, because the glides in these environments — all else being

equal — are predictable. Besides being a predictable glide, /j/ is used to split up non-permitted sequences of three vowels when a vowel initial suffix is added to a long vowel: <asaažumavoq> 'he wants to wash it'. Besides being a predictable glide, /v/ is the first segment of several endings, so there occur morphological inconsistencies, e.g. in the writing of forms like <takuaa> without <v>, as opposed to forms like <nerivaa>, as mentioned in §2. This inconsistency may make children more confused than the problem of the phonetic indeterminacy of approximants alone would cause.

As spelling errors, insertions of predictable glides occur frequently, whereas omissions of morphological approximants occur rarely. In glide insertions the appropriate alternatives most often occur:

- *<sijuani> for <siuani>, [sijwani] 'in the stem of a boat'
- *<siuvani>

But also inappropriate alternatives occur:

- *<qimmiga (<qimmia>), [qimmi^ja] 'his dog'

Glide insertion also occurs within the only diphthong in the language, word final <ai>:

- *<tigupallappagi> (<tigupallappai>) 'he hurried to take them'
- *<ilivaj> (<ilivai>) 'he placed them' (NB. plus omission of final vowel.)

It may be more reasonable to interpret this seeming substitution of *<j> for <i> as a glide insertion with subsequent omission of the final vowel. The postulation of error types like 27 and 28 seems a little ad hoc. However, two statements are true for both: all errors involve a glide/approximant and there is always articulatory similarity between error and target.

Of the substitutions, *<g> for <v> and *<j> for <g> are the most frequent. Rarely they are applied in feeding order: when a child has many substitutions of *<j> for <g> and *<g> for <v>, suddenly a substitution of *<j> for <v> may crop up:

- *<allassavajut> (<allassavavut>) 'the things we are to write',

which is the only occurrence of *<j> for <v> substitution in the data.

A rather odd combination of rules occurs amongst the above-mentioned dialect-speaking children, where /g/: [ɣ] manifests itself as [ŋ]. As a hypercorrection for the dialectal [ŋ] for /g/ they sometimes write *<g> instead of <ng>. They also make the otherwise general substitution of *<j> for <g>. When these two substitutions are applied in feeding order

they yield *<j> for <ng>:

*<tamaviajuat> (<tamaviannguat>) ‘with all one’s might’
(here with wrong quantity)

8. *Vowel substitution.*

Confusion of uvularized/non-uvularized vowels is a general error in all grades. Henceforth the term “vowel substitution” will be used exclusively for this subtype (error type 15 a).

There is a significant decrease in frequency of this error type in Gr. 8 and Gr. 9 (Gr. 9 was only investigated with respect to this error type in order to see if the tendency in Gr. 8 was valid.) Apart from the quantitative difference between the two oldest classes and the rest, there is a very marked difference as to the direction of substitution. Gr. 3 to Gr. 7 have the majority of erroneously uvularized vowels, that is their errors are mainly *<e> for <i> and *<o> for <u>. But in Gr. 8 and Gr. 9 it goes in the opposite direction, and furthermore with an even greater majority, so they make far more errors of *<i> for <e> and *<u> for <o>. It is tempting to suggest that this difference might stem from the older children’s greater awareness of what the underlying phonemes are, serving as a sort of reaction against an inconsistent system. But since both classes started school before the spelling reform, the explanation might be found there. It may be namely an overgeneralization of the repealing of writing <e> and <o> word-finally (Rischel, personal communication). On the other hand, when these two oldest classes do write *<e> and *<o> word-finally, and especially in highly frequent sequences and short words, it seems reasonable to suggest that there it is a direct interference from the old spelling.

A particular subgroup of vowel substitutions is constituted by the contextually adapted vowels in <r>-insertions and <r>-omissions. This will be considered in connection with <r>-errors in §17.

9. <tsi> / <ti>-errors.

Unfortunately this error complex, the insertion and omission of <s> between /t/ and /i/, is unevenly and sparsely represented. The basis for omission of <s>, i.e. the <tsi>-structure, happens only to be represented in the three oldest classes, Gr. 6, 7, and 8, of which Gr. 6 makes no <s>-omissions at all. The basis of inserting <s>, i.e. the <ti>-structure, is rep-

resented in all classes, but only the two oldest, Gr. 7 and Gr. 8, make any <s>-insertions at all. Thus, only Gr. 7 and Gr. 8 make <tsi>/<ti>-errors. Some information is available from comparing these two classes. Both of them have the highest frequency of <s>-omissions as opposed to <s>-insertions. That is, they make relatively more errors like:

*<masatinnagít> (<masatsinnagít>) ‘without making them wet’

and relatively fewer errors like:

*<qitsigissat> (<qitigissat>) ‘dance music’

<s>-omission also occurs sporadically before vowels other than /i/:

*<isikitoq> (<isikitsoq>) ‘small-eyed’

This *may* be taken as a general <s>-confusion caused by the inconsistency in assigning the sibilant only in the long /tt/ before /i/ <nunatsinni> [..tt^si..] ‘in our land, i.e. in Greenland’ but not in the short /t/ before /i/: <atisat> [..t^si..] ‘clothes’. However, there are no occurrences of inserting <s> before vowels other than /i/, so errors like the above-mentioned *<isikitoq> can hardly be taken as caused by a *general* <s>-confusion anyway.

A subgroup within <tsi>/<ti>-errors is that of rendering <tsi> with *<tti>:

*<masattinnagít> (<masatsinnagít>) ‘without making them wet’

This error type can be described as a compensation for the loss of length marking when <s> is omitted. Formally it is a quantity error too whenever <s> is omitted or inserted (cf. §2), but here, as with the other orthographic compromise, the notation of <r>-clusters, a purely formal analysis is of less relevance. In the *<tti>-version, however, length is unmistakably correctly represented, it is only that the wrong alternative has been chosen. On this point there is a marked difference between the two classes. Gr. 8 makes this length-compensated version in 4 out of 5 <s>-omissions, whereas Gr. 7 only makes it in 2 out of 5 <s>-omissions. This fact, in addition to Gr. 8’s rather few <s>-insertions at all in <ti>-structures may suggest that the difference between the two classes arises from Gr. 8’s having started with the old spelling. Gr. 8 has had to learn that old spellings <‘ti>, <vti>, and <gti> all are to be spelled <tsi>. They may therefore have a greater “knowledge” of <s> as a marker of length.

For simplification the examples all showed the <t(s)i> sequence.

Insertion and omission of <s> also occur before the uvularized allophone of /i/, written <e>, e.g.:

- *<atseqarpoq> (<ateqarpoq>) 'he is called'
- *<mingutterpaa> (<mingutserpaa>) 'dirties it'

10. *Quantity errors.*

The three youngest classes are so sparsely represented, and furthermore the teachers here usually go through the dictations in advance, so their error percentages are of less validity. So only the three oldest classes will be considered in detail.

The total frequency of quantity errors falls from 30-40% in Gr. 6 to 5-10% in Gr. 8. Gr. 6 has a higher frequency for vowel errors, Gr. 7 and Gr. 8 for consonant errors. As regards position in the word there is general accordance between where errors *may* occur (i.e. the relevant structures) and where they *do* occur, except that reduction of a geminate consonant occurs less frequently in the first syllable than it does elsewhere in the word. The opposite is true of geminating a single consonant. About 50% of the cases of this error type take place in the first syllable after a short vowel. This is where I suggest a possible interference from Danish length assignment. In Danish long vowels are designated by one consonant following, while short vowels by two consonants following (clusters or geminates). E.g.:

- <hvile> ['vi:lə] 'rest' (The mute <h> being of no relevance for quantity)
- <ville> ['vilə] 'would'

This prosodic/graphemic interference should probably be most likely to occur in short words with short vowels. Words of this type are the only ones comparable to Danish words, and as such subject to interference in storage. The interference may arise from both reading and writing Danish. Of course specific tests are required before I can say anything conclusive.

11. *The correlation between frequency of quantity errors, word length and word complexity.*

In the following paragraph one class (Gr. 7) will be considered in detail.

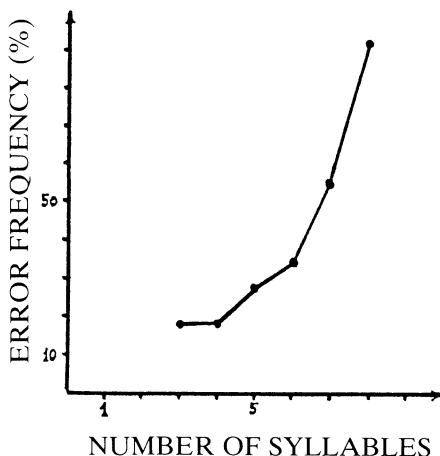


Fig. 1

Fig. 1 shows error frequency (%) for word length. The curve seems to rise more than proportionally. This is true for the other classes as well. In order to get more detailed information it is necessary to separate the two variables, word length and word complexity. Word length will be accounted for by adding up the number of syllables (henceforth S). Concerning word complexity: Kleinschmidt (cf. §2) accounted for length distinction and stress pattern by dividing syllables into four degrees of "weight" (Holtved 1964:29-31,112):

- (C) V : 2
- (C) V C : 3
- (C) V V : 4
- (C) V V C : 5

Assuming that "weight" equals "complexity" it is true that a closed syllable is more complex than an open one, and that a syllable with a long vowel is more complex than one with a short vowel, and hence that (C)VVC is the most complex type of syllable. I assigned the index of complexity by adding up the "weight"-indices of Kleinschmidt (henceforth WK). Some examples:

<erinaq>	(e-ri-naq)	S = 3. WK = 7.
	2+2+3	
<erniinnaq>	(er-niin-naq)	S = 3. WK = 11.
	3 + 5 + 3	

<angajoralimi> (a-nga-jo-ra-li -mi) S = 6. WK = 12.
 $2+2+2+2+2+2$

<marlukkuutaarlutik> (mar-luk-kuu-taar-lutik) S = 6. WK = 20.
 $3+3 +4 +5 +2+3$

Test One: Is complexity the crucial factor for error frequency?

Fig. 2 shows error frequency against WK'-indices for Gr. 7. WK' is WK adjusted for word length, so the curves have a common null point. If complexity is the crucial factor for error frequency the adjusted WK'-curves shall coincide. They do not. Words with a higher value of S score more errors, so word length must be taken into account.

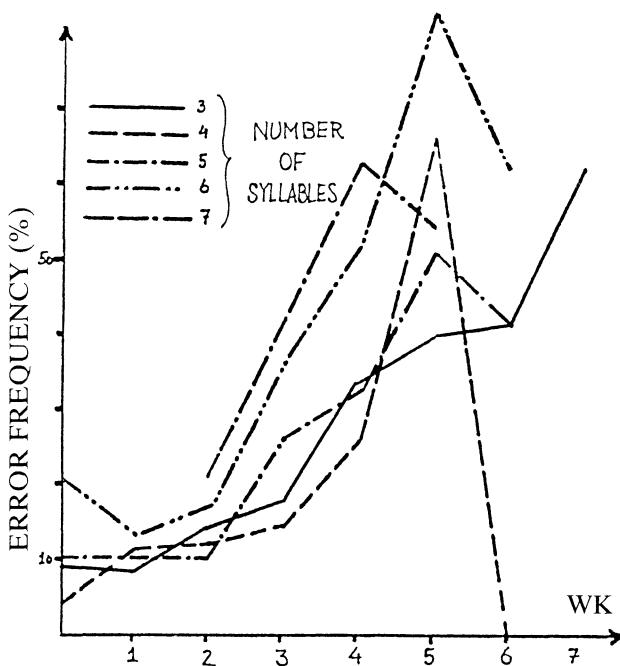


Fig. 2

As can be seen, some of the curves dive at maximum WK'. Since I cannot find any linguistic explanation for this, I suggest that the WK-index invokes some noise. In order to gain purer curves, I tried another index,

namely the number of *graphic geminates* (henceforth G). Some examples: (all three words having six syllables):

- (a) <angajoralimi> : G = 0
- (b) <taamaaqatigipput> : G = 4
- (c) <aallanngitsoorpugut> : G = 4
[aalleŋŋit̪ɔɔ^Kppuŋut] : but phonological geminates = 6

As can be seen from example (c) the number of graphic geminates does not necessarily coincide with the number of phonological geminates, due to the convention of denoting the /r/-clusters and the /ts/-clusters economically (cf. §2). This index considers only graphic geminates as being relevant for complexity.

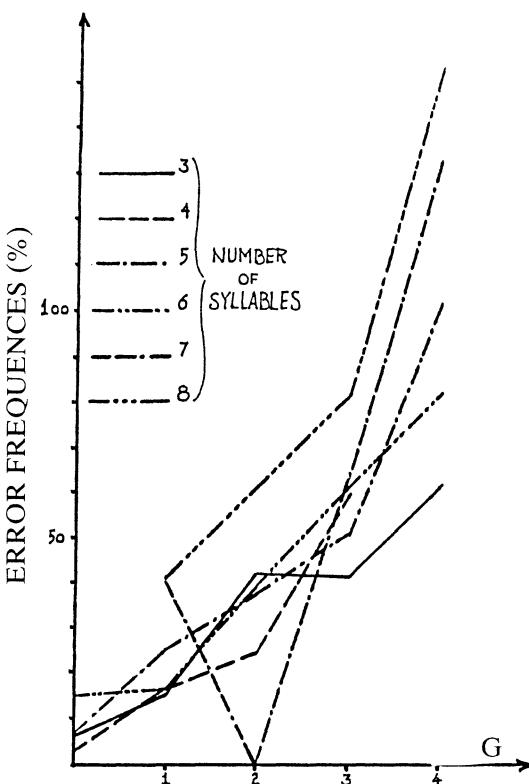


Fig. 3

Fig. 3 shows error frequency against the G-indices. This clearly gives purer curves. They rise more than proportionally. But neither here do the curves coincide, the number of errors with a higher value of S being larger than that with a lower value of S. Hence word length cannot be left out of account.

Test Two: Is word length the crucial factor for error frequency?

Error frequency against increasing S for every value of G separately shows curves which can be taken as proportional, but with constant increase. That is, the greater the value of G the steeper the curves rise. Thus, the direct effect of word length on error frequency is proportional.

But the average number of geminates in words appears to increase more than proportionally. This is probably because suffixation in many cases causes assimilation (vowel and consonant gemination) at the morpheme boundary, or strengthening (consonant gemination) morpheme-internally.

To conclude: the *direct* effect of word length on error frequency is proportional. But as an increase in the number of syllables yields a more than proportionate increase in geminates, and as an increase in the number of graphic geminates yields a more than proportionate increase in the error frequency, the *indirect* effect of word length on error frequency is more than proportional. Fig. 1 shows the multiplied effect.

In Gr. 6 and Gr. 8 there appear similar tendencies. The curves are not identical, since Gr. 6 in general has a higher error frequency and G. 8 a lower one than Gr. 7.

My first intuition was that the *phonological* structure of the word (the WK-index) was the crucial factor for error frequency. So this qualitative variable was separated and tested, and so was the quantitative variable, the word length. It appeared that the plain *graphic* representation, i.e. the G-index, gave purer curves than did the WK-index. Hence, the graphic representation must be a more suitable measure for complexity (error frequency) than the phonological structure.

12. *Multiple quantity errors.*

It is common that words, even short ones, contain more than one quantity error. Two errors in the same word are by far the most frequent of multiple errors. Five are the maximum in my data:

*<issumallioqutanat> (<isumalioqqutaannaat>) ‘only thoughts’

The distance between the errors in multiple errors varies from zero (errors in neighbouring segments):

*<illat> (<ilaat>) ‘some of them’

to a span of several segments:

*<qimmussertalissaput> (<qimussertalissapput>)
‘they are going to make sled-tours’

The distance will be measured in “places”. One place is one error basis, i.e. either one short segment (prone to gemination error) or one long segment (prone to reduction error):

*<m a n n₁s s₁l u> (<manissullu>) ‘even, plane’

 1+1+1 = a distance of 3 places.

In more than 50% of the multiple errors the distance is between zero and 2 places, and a distance greater than 4 places is rare.

It seems plausible that short distance multiple error like *<illat> (<ilaat>) and *<qasussersortarput> (<qassusersortarput>) have more interdependence than errors like *<qimmussertalissaput> (<qimussertalissapput>).

Of words with two quantity errors half the total is constituted by errors going in opposite directions, i.e. a reduction and a gemination occurring within the same word. These errors are labelled ‘bi-directional’. Of the unidirectional errors double geminations occur very rarely. Some typical examples:

two reductions : *<peqinarlunilu> (<peqqinnarlunilu>) ‘being wholesome’

two geminations : *<mammaqqaaq> (<mamaqaaq>) ‘it tastes great!’

bi-directional : *<assapaa> (<asappaa>) ‘he washed it’

*<peqqinaarlunilu> (<peqqinnarlunilu>) ‘being wholesome’

It may also seem plausible that (close) bi-directional errors have more interdependence than any (close) uni-directional errors have.

13. *Some suggestions about cognitive processes.*

Luelsdorff (personal communication) has drawn my attention to the backward displacement of the English consonant doubling rule, and also to the results of Ellis (1979:270-271). Ellis found that the majority of erroneous doublings in his own slips of the pen (perseverative doublings not taken into account) occurred when the word already contained another pair of identical letters, not necessarily adjacent. The doubling errors in Ellis's data are described as caused by erroneous application of letter repetition, e.g. **<innaccurate>* for *<inaccurate>*, triggered by the occurrence of the two *<c>*'s. In Ellis's data the triggering identical pair of letters does not necessarily occur adjacently. But this way of counting triggering identical pairs cannot be applied to Greenlandic, because it has very long words, and only three vowels (and a more limited number of consonants). Consider for example this doubling error:

**<sivikkitsuinnarmik>* for *<sivikitsuinnarmik>*

It would not be possible to decide if it was the two *<s>*'s, the five *<i>*'s (and which of them) or the two *<n>*'s that triggered the erroneously applied doubling of the first *<k>*. The structure in Greenlandic I consider it reasonable to compare with Ellis's data must fulfil the following conditions:

- (1) the triggering identical pair of letters must be adjacent.
- (2) the distance between trigger and error should not be greater than two places (cf. §12).

(Ambiguous cases where there are possible triggers both preceding and following the error are not taken into account.)

The distribution of Greenlandic doubling errors (single occurrences):

For *vowel* doubling errors: (36 in all)

no geminate environments:	25
anticipative doublings	: 11

For *consonant* doubling errors: (102 in all)

no geminate environments:	66
anticipative doublings	: 36

As can be seen the majority of doubling errors occur in words with no other geminate. This is contrary to Ellis's results. The Greenlandic doubling errors thus appear to be of a different kind than the English doubling

errors in Ellis's data.

But it is interesting that Ellis's interpretation of doubling errors as erroneous application of features is of relevance to the otherwise incompatible bi-directional quantity errors (cf. Luelsdorff 1986). I consider them incompatible with errors like *<innaccurate> because bi-directional errors *eo ipso* invoke what could be called mutual compensation. Ellis does not report on such mutually dependent errors, (e.g.: *<innacurate>). But still the expression 'displacement of feature' is relevant.

Another device useful in the description of certain Greenlandic quantity errors is the suggestion of Wing and Baddeley of a hierarchical representation of words stored in the buffer (Wing & Baddeley 1980:272).

The error frequency of bi-directional errors decreases from 77% in Gr. 3 to 30% in Gr. 8. This marked change suggests a change in spelling strategy in the development from unskilled to skilled speller. Supposing a hierarchical storage for young spellers but not for the oldest this can explain their different error patterns. For unskilled spellers prosody is represented separately as characterizing the word as a whole, but not as features of length for each phonemes. (Notice that quantity, stress and intonation are interrelated in a complex way, cf. Rischel 1974: 91 ff). A bi-directional quantity error can then be interpreted as a failure to transfer prosodic patterns into features of [\pm GEMINATE]. This can be called 'displacement of features'.

The young spellers do have a certain knowledge of the relation between stress pattern and the distribution of single and geminate letters, but they are so to speak caught between two layers of information. Only gradually do they master specifying the features of length for every phoneme in particular. Then storage is no longer hierarchical, and hence very few bi-directional quantity errors occur.

The error frequency of double reductions, however, goes in the opposite direction: it increases from 15% in Gr. 3 to 70% in Gr. 8. (Note that the numbers are % of multiple errors in total. In general the quantity-error frequency decreases from Gr. 3 to Gr. 8 (cf. §10)). At a first glance it looks as if the two developments are mutually dependent, the fall of one being related to the rise of the other, so that it is their combination that expresses a certain shift in spelling strategy. But probably the latter development does not express any shift in spelling strategy. Double reductions occur mainly in long words, and as word length also increases from Gr. 3 to Gr. 8, double reductions may be considered as two occurrences of single reduc-

tion in words beyond a certain length. This issue requires a specific investigation of the size of programming units in the spelling process (cf. e.g. Nauclér 1980: 142 ff.) Such an investigation is absolutely necessary for the study of writing a polysynthetic language like Greenlandic. Further, an account of the size of programming units and an adequate description of multiple quantity errors are closely interdependent.

14. *Immediate environments of quantity errors.*

For the two oldest classes I have mapped out the immediate environments of reduction errors (cf. Wing & Baddeley 1980:271) to see if there were different error frequencies in different environments. The relevant range here being 4 places, i.e. the four immediately surrounding vocalic and consonantal segments, whether short or long. (^{t/r} being an abbreviation for /r/-clusters and /ts/-clusters in common.)

(1) *Error frequency before and after specific structures.*

Gemинates.

Both classes have a greater error frequency *after* geminates (whether CC or VV) than before

t/r-clusters.

But as regards the two clusters they differ, Gr. 7 having greater error frequency *before* clusters and Gr. 8 *after*. In spite of this different status of clusters in the two classes it is still true also for Gr. 8 that the G-index gives purer curves than does the WK-index (cf. §11), which supports the relevance of the geminate-based measurement of complexity.

(2) *Error frequency in single versus geminate environment.*

For both classes the error frequency is far greater after any *geminate* than after any single segment.

(3) *Crucial environments.*

Finally, the differences in crucial environments are — other things being equal — more significant as regards *preceding* configurations than following ones.

15. *<r>-insertions.*

The vast majority of *<r>-insertions* take place before *<q>*, (250 instances out of 319). As exemplified in §2 the *<r>* in /r/-clusters has a double function: it denotes both the uvularity and the long consonant. When inserted in the last syllable it is evidently a uvularity-marker: **<naterq>* (*<nateq>*) 'floor' as it is in **<utarqqivarput>* (*<utaqqivarput>*) 'we were waiting for him' and in **<torqqusimavoq>* (*<toqusimavoq>*) 'he is dead', where a quantity error is manifested in the doubled **<qq>*. But in errors like **<torqusimavoq>* (*<toqusimavoq>*) it is not a priori clear whether this error is a quantity error. (The old spelling wrote *<rq>* for /qq/). The question is further complicated in that there is only one *<a>*-grapheme, so the sequence **<arq>* (*<aq>*) does not yield redundancy as regards uvularity as a sequence like **<orq>* (*<oq>*) does. It appears that there is no preference for inserting *<r>* after specific vowels. Indeed it would have simplified the matter if there had been such a preference, were it after *<a>* or after *<e/o>*.

Now, *<r>* is also inserted before consonants other than *<q>*, (69 instances out of 319). There appears no clear phonetic pattern in this subtype. I therefore investigated if this kind of *<r>-insertion* could be explained in terms of the property of *<r>* as a length marker, i.e. if it is inserted to account for a CC-structure. In order to get an unambiguous picture the following words were dismissed:

- (1) all instances where **<k>* is substituted for *<q>*, since this is a general substitution error (cf. §6)
- (2) all instances with a plausible trigger:

**<arlarsuaq>* (*<allarsuaq>*) 'great land bear'

- (3) all instances with simultaneous insertion and omission of *<r>*:
- **<arlasuaq>* (*<allarsuaq>*)

This reduced the inventory to 30. Out of these there were 14 instances of **<rC>* for /CC/. Before suggestions can be proposed as to an active use of *<r>* as a length-marker it is necessary to look at the whole error pattern of the individual child. Especially one must take into consideration how the child makes

- (1) vowel substitutions
- (2) quantity errors (reductions and geminations)

- (3) <r>-omissions
- (4) other types of <r>-insertions

Ad (1): for the role of concomitant vowel substitutions in <r>-errors see §17.

Ad (2): an active use of <r> as a length-marker is an advanced strategy, a subtlety which cannot be postulated if the child is a poor quantity speller.

Ad (3) and (4): if the child is generally “<r>-confused” and makes many other insertions and omissions of <r>, this is strong evidence against the proposed subtlety. If, however, the child makes almost no other types of <r>-insertions, and his/her <r>-omissions are mostly length-compensated (cf. §16), this is consistent with the assumption about use of <r> as a length-marker. Only 5 out of the 14 possibilities satisfied these criteria:

- *<milerpoq> (<milippoq>) ‘it is blocked up’
- *<isserpoq> (<issippoq>) ‘the weather is cold’
- *<atuammersortiarluni> (<atuammersutsiarluni>) ‘reading for quite a while’
- *
- ” ” ”
- *<atuammersurtiarluni> ” ”

Notice the last three words at the same time count as length-compensated <s>-omissions (cf. §9). (They were produced by three different children.)

To conclude: since the vast majority of <r>-insertions take place before <q>, and since only 5 cases (of 14 possible) of active use of <r> as a length-marker could be discerned, and since only 3 cases of 147 <r>-insertions between <e/o> and <q> had concomitant de-uvularization of the vowel (*<irq>, *<urq> for <eq>, <oq>), which is below probability for substitutions in general, it seems reasonable to describe <r>-insertions as being mainly a homogeneous error type, namely an explicit and redundant marking of the vowel’s uvularity.

I still have not accounted for the <r>-insertions before consonants other than <q>. Some of these (the one weeded out above) apparently have a contextual trigger. But there still remains the odd ones for which I have no suggestion thus far.

16. <r>-omissions.

Neither do <r>-omissions occur preferentially after specific vowels. Depending on vowel adjustment and/or length compensation there are four ways of *not* writing <r> in front of a consonant:

For <iserpoq> 'he went in'

- (1) *<isepoq>
- (2) *<iseppoq>
- (3) *<isippoq>
- (4) *<isipoq>

As the uvular vowel indicates a uvular consonant following, and consonant length is predictable in /rC/-clusters, (1) contains sufficient information as to the structure of [^KCC] and it is highly non-redundant. But it breaks the rule of indicating long phonemes by two segments. (2) is redundant as regards consonant length. But where (1) must be labelled as a formal quantity error, (2) unambiguously accounts for the long consonant, and will therefore be labelled 'length-compensated <r>-omission' (cf. also the length-compensated <ts>-errors in §9). Though formally a gemination error it is evidently not a quantity error. (3) and (4) on the other hand are serious errors, both of them missing the uvular element, and (4) even the long consonant. The omission of uvularity is a crucial factor since <isippoq> actually is another lexeme, meaning 'he fell into the water from the shore'. The ending <-poq> is morphologically inherent to consonant stems, where the assimilation process produces <pp>. (Vowel stems always have <-voq>-endings.) Since the double <pp> in this particular case is predictable, (4) too would render <isippoq> unambiguously.

If one were to suggest some guidelines as to differentiation between the four error types when giving marks, it must clearly be judged unreasonable to expect teachers to distinguish four degrees of error and give four different marks. One might estimate (1) and (2) as equal and give each of them one minus point, and similarly give (3) and (4) two minus points each.

But a differentiation of four degrees is not possible at all in words with <a> preceding the omitted <r>:

For <nalunaarpoq> 'he announced'

- (5) *<nalunaapoq>
- (6) *<nalunaappoq>

Actually (6) is another lexeme, meaning 'it is distinct'. But whereas (4) *<isippoq> got more minus points than did (2) *<iseppoq>, because it unambiguously denoted another lexeme, it seems on the contrary reasonable to give (6) the highest score (fewer minus points) because of the length compensation which denotes the length correctly. And, whereas (1) *<isepoq> and (2) *<iseppoq> obtained equal scores, this is unreasonable with (5) and (6), because <p> versus <pp> is precisely the only difference

between them. So the scope of <r>-omission is quite different depending on the preceding vowel, and it is not possible to carry out a homogenous description of this error type.

As regards the length-compensated <r>-omissions in total, that is all errors like (2), (3) and (6) above, the frequency increases from 5% in Gr. 4 to 55% in Gr. 8. This might suggest an increasing ability to detect that length is insufficiently indicated when <r> is missing, or in other words that the psychological reality of <r> as a marker of length is increasing. This coincides with the increasing proficiency in denoting quantity in general, as mentioned in §10. It also coincides with the increasing proficiency in assigning the feature [+ long] to specific phonemes, (cf. the suggestions in §12 and §13). The length-compensated <r>-omissions unambiguously show that the second member of the /rC/-cluster is long. (/rC/ is manifested as [^bCC]).

The increasing frequency of length-compensated <r>-omissions is also true for <e/o>-words separately. Here it is accompanied by a steep decrease in the frequency of erroneous de-uvularization of the preceding vowel, from 100% in Gr. 4 to 0% in Gr. 8. That is, from Gr. 4 to Gr. 8 there occur still fewer errors like

*<inneq> (<erneq>) 'son'

in favour of still more errors like

*<enneq> (<erneq>)

This might suggest the increasing psychological reality of <r> as a *redundant* marker of *uvularity*, which coincides with the status of <r> in <r>-insertions (cf. §15).

To conclude: the increasing frequency of length-compensated <r>-omissions together with the decreasing frequency of concomitant erroneous de-uvularization of the preceding vowel show that, in the development from unskilled to skilled speller, the *double* function of <r> as a *length*-marker for consonants and a *uvularity*-marker for vowels becomes more established as a psychological reality.

17. *Vowel adjustment in <r>-errors.*

Most often <r>-omissions and <r>-insertions before a consonant other than <q> are accompanied by contextual adjustment of the preceding vowel:

<r>-omission with vowel adjustment (de-uvularization):

*<atisalisupunga> (<atisalorsorpunga>) 'I put on my clothes'

<r>-insertions with vowel adjustment (uvularization):

*<isserpoq> (<issippoq>) 'the weather is cold'

This might suggest that the two types of sequences, <e/o> + <r/q> and <i/u> + C are rather well established as grapho-motoric stereotypes. Since vowel substitution is a general error, it is not a priori evident what the "core" of these combinations is. Either it is (1) a vowel substitution which triggers a contextually appropriate consonant or (2) an <r>-error which triggers a contextually appropriate vowel.

Assuming the former alternative in <r>-insertions like *<isserpoq> (<issippoq>), it is extremely striking that it is exclusively <r> which is triggered and never <q>, since <q> is just as motivated after a uvular vowel as <r> is. The resulting erroneous word thus always contains a permitted sequence, <rC>, whereas forms like *<isseqpoq> with the unpermitted sequence *<qC> do not occur.

As regard <r>-omissions, the former alternative can only account for the first two <r>-omissions, but not for the last one (after <a>) in:

*<qassusisutaput> (<qassusersortarput>) 'they throw out nets'

Recall also that neither <r>-omissions nor <r>-insertions occur preferentially after specific vowels (cf. §15 and §16).

So it seems reasonable to suggest the latter alternative and consider <r> to be the core of these error combinations. The vowel adjustment is thus secondary, caused by established grapho-motoric stereotypes.

Concluding remarks.

The most prominent error types of Greenlandic school children are quantity errors, /rC/-cluster errors and uvular/non-uvular vowel substitutions. These vowel substitutions arise from the orthography's overdifferentiating vowel inventory, with five vowel letters accounting for three vowel phonemes. The /rC/-cluster errors are to a certain extent caused by the orthographic compromise inherent to the principle of graphic economy.

Quantity errors are prominent in languages with quantity distinction. It is important to find out which error patterns are general and which are language specific. In Greenlandic the multiple quantity errors are of special interest for further investigation.

ACKNOWLEDGEMENTS

I want to thank the *teachers* for their kindness and collaboration during my collection of data, and Professor Robert Petersen (Ilisimatusarfik, Inuit Institute, Nuuk) and Grønlands Hjemmestyre for moral support.

I also want to thank Professor Dr. Jørgen Rischel (Institute of Phonetics, University of Copenhagen) and Lecturer Frans Gregersen (Institute of Nordic Philology, University of Copenhagen) because they always showed interest in my project, and always had time for discussion, encouragement and good advice, mag. scient. Peter Colding-Jørgensen for helping me with the mathematical evaluation of my data, Lecturer Michael Fortescue (Institute of Eskimology, University of Copenhagen) for weeding out the worst Danicisms and other linguistic errors, and external lecturer Christian Berthelsen (Institute of Eskimology, University of Copenhagen) for correcting my translations of the Greenlandic examples.

Finally, I want to thank Dr. Philip A. Luedsdorff (Institute of English, University of Regensburg) for much help and inspiration during preparation of the manuscript.

NOTE

1. Further studies have revealed that the majority of vowel substitutions are of the type: *<i>/*<u> for <e>/<o>. This suggests that the children generally have a strong sense of <i> and <u> being the proper phonemes.

BIBLIOGRAPHY

- Ellis, Andrew W. 1979. "Slips of the pen," *Visible language*, 13: 265-282.
 Frith, Uta. 1980. "Unexpected spelling problems." In: Uta Frith (Ed.) *Cognitive processes in spelling*. London: Academic Press, 495-515.
 Holtved, Erik (Ed.). 1964. *Kleinschmidts Briefe an Theodor Bourquin*. Meddelelser om Grønland. Vol. 140, No. 3. København: Reisel.
 Luedsdorff, Philip A. 1986. *Constraints on error variables in grammar: Bilingual misspelling orthographies*. Amsterdam: John Benjamins.
 Mase, Hideo & Jørgen Rischel. 1971. "A study of consonant quantity in West Greenlandic," *Annual report of the Institute of Phonetics, Univer-*

- sity of Copenhagen*, 5:175-247.
- Moats, Louisa Cook. 1983. "A comparison of the spelling errors of older dyslexic and second-grade normal children," *Annals of dyslexia*. 33:121-139.
- Nauclér, Kerstin. 1980. *Perspectives on misspellings. A phonetic, phonological and psycholinguistic study*. Malmö: Gotab.
- Nelson, Hazel E. 1980. "Analysis of spelling errors in normal and dyslexic children." In: Uta Frith (Ed.) *Cognitive processes in spelling*. London: Academic Press, 475-493.
- Petersen, Robert. 1970. *Rapport: Retskrivningsprøver*. Mimeogr. Ed. by The Ministry of Greenland, København.
- Rischel, Jørgen. 1972. *Bemærkninger til det af Landsrådets sprog- og retskrivningsudvalg udarbejdede forslag til en ny grønlandsk retskrivning*. Unpublished ms. København.
- Rischel, Jørgen. 1974. *Topics in West Greenlandic phonology*. København: Akademisk Forlag.
- Wing, Alan M. & Alan D. Baddeley. 1980. "Spelling errors in handwriting: A corpus and a distributional analysis." In: Uta Frith (Ed.) *Cognitive processes in spelling*. London: Academic Press, 251-285.

THE ORTHOGRAPHIC REPRESENTATION OF NASAL VOWELS IN ACEHNESE¹

Mark Durie

Acehnese is a language of north Sumatra with a set of nasal vowel phones in contrasting distribution with oral vowels. In a classical phonemic analysis Acehnese would have a separate set of nasal vowel phonemes. Due to factors described in this paper, the spelling of Acehnese is not standardised; it is largely an individual matter how one spells. Despite this lack of a prescriptive standard it is a striking fact that in no orthography of the several that are in use or have been suggested are the nasal vowels consistently treated as independent phonemes. Instead all orthographies have certain features in common which suggest a non-phonemic phonological representation of vowel nasality. It would seem that where vowel nasalisation is concerned the orthographies do not encode phonemic phonological structure, but an at least partially unsegmented phonology where vowel nasalisation is not marked as a feature of one segment, but is located *prosodically* over a larger unit of structure than the segment. Moreover this representation reflects markedness: only the marked cases are specifically distinguished. The conscious intuitions of native speakers, as reflected in their orthographies, appear to access this prosodic level of representation, including the markedness, more directly than the segmental phonemic level. I use the term *level* of phonological representation to suggest that speakers have a multi-systemic phonology, with coexistent representations that are organised rather differently. The orthographies themselves are segmental, since they are composed of a linearly ordered arrangement of letters for vowels and consonants. This means that they can only represent a prosodic phonology indirectly.

The Phonology of Acehnese

The vowels of Acehnese (using IPA symbols) are:

Monophthongs

ORAL

high	i	w	u
high-mid	e	ɤ	o
low-mid	ɛ	ʌ	ɔ
low		a	

NASAL

high	ī	ū	ū
mid	ɛ̄	᷃	᷅
low		᷁	

Diphthongs

ORAL

high	i ^ə	w ^ə	u ^ə
mid	ɛ ^ə		ɔ ^ə

NASAL

high	ī ^ə	ū ^ə	ū ^ə
mid	ɛ̄ ^ə		᷅ ^ə

For detailed information on the articulation of these vowels see Durie 1985.

The consonants are (using IPA, except that y = IPA j, and j = IPA ɻ):

	<i>Consonants</i>									
	<i>bilabial</i>	<i>alveolar</i>	<i>palatal</i>	<i>velar</i>	<i>glottal</i>					
<i>oral obstruents</i>										
voiceless										
stops	p	p^h	t	t^h	c	c^h	k	k^h		
voiced	b	b^h	d	d^h	j	j^h	g	g^h		
trills			r	r^h						
laterals			l	l^h						
fricatives			s		f					
<i>glides</i>										
continuants	w			y			h			
stop							?			
<i>nasals</i>	m	n		p		ŋ				

The glides /w/, /y/, /h/, /?/ have nasalised allophones in the environment of a nasal vowel in the same syllable. Thus /sāh/ [sāħ] ‘whisper’ and /mūħāy/ [mūħāŷ] ‘expensive’.

For the purpose of this discussion it is helpful to distinguish the following four types of consonant + vowel sequences:

- CV - oral consonant + oral vowel: /lu^{2i²}/ ‘muddy, slippery’; /sa^h/ ‘legal’; /baŋāy/ ‘stupid’; /māta/ ‘eye’
- N᪔ - nasal stop + nasal vowel: /baŋāy/ ‘stupid’; /mān⁵/ ‘bathe’; /nāŋ/ ‘relative clause introducer’; /māta/ ‘eye’
- C᪔ - oral consonant + nasal vowel: /sāh/ ‘whisper’; /cahēt/ ‘sever with a knife on the end of a pole’; /nāp^hāh/ ‘breath’
- NV - nasal stop + oral vowel: /ŋay/ ‘type of bird’; /tiŋay/ ‘remain, leave’

Word stress falls on the final syllable; this syllable can have all four sequence types given above. In unstressed (non-final) syllables only the types CV and N᪔ occur, so /māta/ ‘eye’ and /puta/ ‘turn’ are possible, but not /mata/ and /pūta/. Of the four sequence types, CV is the most frequent, both in stressed and unstressed syllables. N᪔ is also very frequent. Com-

pared to these two C^V and NV are much less frequent, NV being rather more common than C^V.

In orthographies the full phonemic status of nasal vowels is obscured in the following way:

/pa/ would always be represented as *pa*

/mā/ would always be represented as *ma* (never marked nasal)

/pā/ would typically be represented as *pa* or *p'a* (also: *pə*, *Pa*)

/ma/ would typically be represented as *mba* (also: *ma*)

Thus a word like /mon/ ‘dew’ would typically be written *mbōn*; /mɔ̄/ ‘cry’ as *moe*, with no mark of nasalisation; and /p̄īp/ ‘suck’ and /t̄ān/ ‘until’ as *piep* or *p'iep* and *an* or *'an* respectively.

Literacy in Acehnese

Literacy in Acehnese is quite a different affair for an Acehnese person from literacy in English for an English speaking person, both in its acquisition and application. The first stage in the traditional formal education of an Acehnese person is to attend classes in the village where a group of young children learn to recite the Koran. The word for ‘read’ in Acehnese is *beuet* /būt̄/ and refers primarily to the recitation of the Koran. Children at the village class learn to pronounce the Arabic alphabet and particularly to recite a few short important passages from the Koran. The meaning of the passages is explained to them, but they do not learn to understand the Arabic directly: the important task is to be able to recite correctly. The ability to recite the Koran, even if only a few passages, is highly valued, and very special emphasis is placed on the phonetic accuracy of the recitation. The teacher who imparts this skill to children is spoken of as making the children into people (*peujeuet keu ureueng*). The recitation is melodic, the rises and falls and the timing of the melody are strictly determined by traditional practice. Public Koran reciting competitions are immensely popular in Aceh today; thousands of people may gather in the open air to listen. The under-age divisions are particularly enjoyed since the spectators can marvel at the skills of the young competitors. For most Acehnese it is not possible to acquire more than the ability to recite a few passages.

The next stage in the traditional acquisition of literacy involves learning *basa jawoe*, ‘jawoe language’, that is, Malay written with an adapted Arabic script. This is the usual language of written prose in traditional Acehnese society. Religious textbooks are for the most part in Malay and letters and other documents such as passports, laws and seals of authority

were traditionally in Malay. Only a minority achieved a proficiency in this written Malay. For a young man, learning it would mean attending a *pesantren*, a religious school, for a period of perhaps two or three years. The emphasis in learning written Malay is on the ability to read since few people had much opportunity for writing, except in copying texts already committed to paper.

Via an ability to read *basa jawoe* an Acehnese person could acquire the skill of reciting Acehnese from the traditional Arabic derived script. This uses the Arabic script with the five additional letters for writing Malay (and Acehnese) to represent /c/, /ŋ/, /p/, /g/ and /j/: these phonemes do not occur in Arabic. With very rare exceptions, written Acehnese is in verse, in a form called *sanjak* /saŋja²l/, which is chanted using a melody. In most cases the person reciting the text would be familiar with its contents, and the script serves as an aid to the memory. Sometimes such texts are in fact committed to memory and recited without the aid of a written text. It is not necessary that the orthography should give an accurate representation of all words because the reciter knows the contents fairly well, and the rhyme gives considerable help in interpreting the written text. The verse is organised into lines with, in its fullest form, two internal rhymes, and a final syllable which rhymes with the surrounding lines. Thus each line contains five rhyming syllables:

The final rhyme, here /-a/, is typically maintained over hundreds of lines. The second internal rhyme, here /-e²/, may also be maintained over many verses; in this particular poem (*Hikayat Jugi Tapa*) it was continued for 98 lines. The first internal rhyme varies freely from line to line.

The form of verse just illustrated is unusual in its complexity. More common is a form with 16 instead of 20 feet per line. The first internal rhyme is omitted in this simpler form, and the second rhyme varies freely from line to line. The line-final end rhyme remains the same as for the more complex form: it is invariant over many verses. An illustrative passage from the same poem as the above is:

nī kubayan / nūk^hun mūnō³ // ta²e² kunō³ / bijeh māta
 sinō³ tādu³ // sajan ulon // pōcut ampon / sajan hama
 tapi 2anū² / bē² t^hat riyoh // lon takot troh / jugi tapa

It should be obvious that the form of verse, especially with the 20 foot line, but also with the more common 16 foot line, is a great aid to reciting.

When a reciter comes to a phrase which he or she cannot decipher and cannot remember, something is made up on the spur of the moment which will fit. It is important in public recitation not to interrupt the flow of the melody. There are certain fixed phrases such as /he se³dara/ 'hey relatives' and /he 2ade²/ 'hey (younger) brothers' (see the first extract above) which can often be inserted into such a gap.

For these three reasons — the familiarity of the text, the help given by the rhyme, and the possibility of making something else up — it is not necessary for the script to represent all the contrasts of the spoken language. Indeed the traditional script is far from phonetically accurate. To represent all the vowels of Acehnese, the Arabic script has only three vowel letters, for *i*, *u* and *a*. A further complication is that the background of literacy in Malay which an ability to recite Acehnese presupposes means that often if there is a similarly sounding Malay word the Malay spelling is simply used for the Acehnese word. Thus the relative clause introducer /pāŋ/ is typically written ي like the Malay *yan*. Final -*s*, -*r*, and -*l* of similar Malay words often appear in the Acehnese script: تيكوس *t.i.k.u.s* for /tikoh/ 'mouse' (Malay *tikus*); اولر *2.u.l.r* for /2ulur³/ 'snake' (Malay *2ular*); کاپل *k.a.p.l* for /kapay/ 'ship' (Malay *kapal*). Often Acehnese /ɔ³/ corresponds to Malay *i*, /ɛ³/ to Malay *u* and /u³/ to Malay *a*, so: جاري *j.a.r.i* for /jarɔ³/ 'hand' (Malay *jari* 'finger'); گورو *g.u.r.u* for /gure³/ 'teacher' (Malay *guru*); and /2ulur³/ given above. There are several other types of Malayisms: کودا *k.u.d.a* for /guda/ 'horse' (Malay *kuda*); سمبایان *s.m.b.h.i.y* for /sumāyan/ 'worship' (Malay *sēmbahyan*); فنچان *p.n.j.y* for /panjāy/ 'tall, long' (Malay *panjāy*). The extent to which writers use such Malayisms varies in its details, but it is always quite substantial. A further complication is that words which have been borrowed into Acehnese and adapted to its phonology are often rendered in their Arabic form in the script, although they are intended to be recited as Acehnese words. The degree to which this is done depends of course on the writer's knowledge of Arabic. Two examples would be هادر *h.a.d.r* (*ha:dir*) for /hale/ 'present' and قربان *q.r.b.a.n* (*qurba:n*) for /kurubuw³n/ 'sacrifice'.

Of those people who can recite Acehnese, few would ever have occasion to write anything original; their own writing would be restricted to copying manuscripts. This means that few people need to decide how to spell words and if an individual feels the urge to creativity, either to write an original work or to modify a work being copied, then he or she has to work out how to spell unusual words using his or her own discretion. There is no prescriptive standard. It is clear why no such standard has developed. In the first place Malay had always been the language of the courts in the politically dominant port kingdoms: Acehnese literature developed in localities which were not politically in a position to serve as centres from which a standard could be disseminated throughout the whole of Aceh. Moreover Acehnese dialects vary considerably; spelling tends to follow local practice. Another important reason is that it is difficult for anyone to read more than a few manuscripts. They are regarded as very precious and their owners are usually very reluctant to lend them, even to friends and relatives. Often the texts have a religious or occult value and their free circulation would be seen as diminishing their value. Collecting manuscripts is very difficult: a collection of as many as a dozen texts might take years to acquire. This means that any one person typically only has a limited experience in reading Acehnese. Divergent spelling practices can very easily develop, and writers are of necessity very dependent on their own intuitions and often unreliable knowledge of Malay and Arabic spelling.

The Jawoe Orthography

In the script derived from Arabic the contrast in vowel nasalisation is marked after nasal stops by interposing an oral voiced stop before oral vowels. The contrast is marked after /ʔ/ by writing the letter 'ain ئ' as the consonant letter before nasal vowels. Before oral vowels *alif ئ* or *hamza ؤ* is used to represent the consonant /ʔ/. In the environment after all other consonants the nasal: oral contrast is not marked in the orthography. Thus the four consonant + vowel sequence types described above are represented as follows:

CV is written as consonant + vowel, or in accordance with the Malay orthographic convention, simply as a consonant, with the vowel understood. Thus *تُوْلِي* *t.u.l.η* for /tulw³η/ 'bone'.

NV is represented just like CV: *مُنِيْ* *m.n.i* for /mūn³/ 'like this'.

CNV is written just like CV, with no marker of nasality, except when the

consonant is /?²/; قَبَ p.i.b for /pī²p/ ‘suck’; but compare ظَعَفَ .u.h for /?²h/ ‘until’ with ظَمِيلَ i.m for /?²i²m/ ‘be quiet’.

NV is represented as nasal stop + homorganic voiced stop (+ vowel): خَنْخَةَ η.g.η for /ŋan/ ‘type of bird’. Not uncommonly this voiced stop letter is left out, giving a spelling indistinguishable from that for NV. Sometimes a similar sounding Malay word will cause a voiced stop in the spelling where Acehnese has a nasal vowel, so /mānə^e/ ‘bathe’ is spelled sometimes like Malay *mandi*; more often in such cases Acehnese will have an oral vowel, as for /tiŋay/ ‘dwell’ which is spelled like Malay *tinggal*.

Modern Roman Orthographies

During the last thirty or so years many publications have appeared in Aceh of Acehnese traditional literature. Some of these use the *Jawoe* orthography described above, and a few use a variant of Snouck Hurgronje’s orthography, described below, but many use an idiosyncratic romanisation of the *Jawoe* script. These publications come in a wide variety of formats. Many are produced from a typed stencil. Writers of these texts attempt to regularise phonetic correspondences when transferring from the Arabic to the Roman orthography. However as far as vowel nasalisation is concerned the modern orthographies used in publications typically use essentially the same system of contrasts as the traditional script. Instead of the Arabic letter ‘ain they mostly use an apostrophe, and a sequence ?V, with an oral vowel, is simply written V in word initial position, and sometimes with a hyphen for the /?²/ in word internal position, thus: *eh* for /?²eh/ ‘sleep’; *la-en* for /la?²en/ ‘other, different’; but *’oh* for /?²h/ ‘until’.

Snouck Hurgronje’s Orthography

In his pioneering work on Acehnese phonology Snouck Hurgronje (1892) developed a Romanised orthography which represented the sequence CV as a nasalised oral consonant (sic) + a vowel. The nasalisation of the consonant was to be indicated orthographically by a different type face: in practice usually italics or bold. Thus, using capitals here for the nasalised consonants: *Piep* for /pī²p/ ‘suck’; *Han* for /hān/ ‘not’. Although it is clear from Snouck Hurgronje’s 1892 article that he was aware of the nasalisation of the following vowels, he seems to have regarded these nasalised consonants as a phonetic reality. Perhaps he was influenced by

the traditional use of Arabic ‘ain for /ə/ before a nasal vowel. Professor Drewes informs me that in Arabic ‘ain conditions nasalisation of the following vowel (personal communication). Indeed Acehnese borrowings from Arabic preserve this nasalisation: /dɔ²ā/ ‘pray’ from Arabic *du'a:*. Snouck Hurgronje rendered NV in his orthography as NCV, with an intervening voiced stop. He was aware that this stop combination, as he regarded it, of nasal + voiced stop, was pronounced as a single segment, but it seems from what he wrote that he thought of this is a kind of allophony, whereby the oral stop is assimilated to the preceding nasal.

Budiman Sulaiman's Orthography

In his studies of the Acehnese language the Acehnese scholar and native speaker Budiman Sulaiman developed an orthography which is described in his grammatical description of 1977. His innovation is to recognize the independent status of nasal vowels after oral consonants, which he marks by a preceding apostrophe. Thus *h'an* for /hān/ ‘not’; *kh'ēb* for /k^hēp/ ‘bad smell’. He does not identify these vowels with the nasal vowels after nasal stops. He marks the contrast NV: NV in the traditional way by inserting a homoorganic voiced stop before an oral vowel, thus *ma* for /mā/ and *mba* for /ma/.

Abdul Gani Asyik's Phonemic Analysis

Another native speaker of Acehnese who has described Acehnese phonology is Abdul Gani Asyik, who provides a phonemic analysis in his 1972 Master's thesis. Like Budiman Sulaiman he recognises the full phonemic status of nasal vowels after oral consonants, but he also fails to identify these with the nasal vowels after nasal stops. He differs from Budiman Sulaiman in suggesting a phonemic contrast between plain nasal stops (with following nasalized vowel allophones) and ‘incomplete’ nasal stops, which he distinguishes with an underlining. Thus his /ma/ for my /mā/ and his /ma/ for my /ma/. Abdul Gani Asyik's analysis is very accurate in other respects, so it is striking that he contrasts NV and NV sequences in terms of their consonants. He mentions vowel nasalisation, but only as allophony conditioned by the preceding plain nasal stop. He does not realise that these ‘nasalised vowels’ are in fact identical with the independently nasal vowel phonemes. The nasal stops before oral vowels, which Abdul Gani Asyik analyses as incomplete nasals, are somewhat longer in duration than

nasals before nasal vowels, and they have slightly less nasal air flow (Durie 1985:15), but by far the most prominent phonetic difference is the vowel nasalisation, and the set of nasal vowel phones following nasal stops corresponds exactly to that of the nasal vowels following oral consonants. By the criteria of phonemic analysis it is clear that this should be regarded as a vowel contrast.

Evidence of Informants' Intuitions

Whilst doing field work on Acehnese, I found that informants easily recognised the nasal vowels occurring after oral consonants as distinct sounds, and could provide further examples of such sequences when I asked for them. These examples would not include instances of nasal vowels after nasal stops. The contrast of NV with N̄V was perceived, as far as my method of questioning could reveal, as a consonant contrast. This seemed to be the case for literate and illiterate speakers alike. Informants grouped examples illustrating this contrast just as if it was a consonant contrast. An informant might volunteer /pəŋ/ 'sugar cane crusher': /pəñ/ 'yes' and /saja²/ 'verse type': /majā²/ 'infant' as examples of the same contrast between two types of /p/. Similarly /hama/ 'slave': /hamā/ 'chicken louse' and himē³ /'type of ape': /tumē³/ 'get' might be offered together as examples of a contrast between two types of /m/. However a pair such as /hān/ 'not': /tuhan/ 'god' would never be grouped with /bunā/ 'genuine': /buña/ 'thing, valuable article' as showing the same contrast.

Rhyme

In Acehnese poetry rhyme is borne on the stressed syllables, and depends on the similarity of the vowel and any final consonant of this syllable. /lu³ŋ/ 'ditch' rhymes with /nūtū³ŋ/ 'you receive' and /jars³/ 'hand' with /mūntrɔ³/ 'minister of government'. Not uncommonly the rhyme is only partially accurate, so /lon/ 'I' may be rhymed with /gampɔŋ/ 'village', or /bri/ 'give' with /ile/ 'flow'. An investigation of rhymes involving nasal vowels shows that these are distinguished from oral vowels. Thus /sumāwē/ 'place name' rhymes with /ramē/ 'crowded, many'; /mūdo²ā/ 'pray' with /tumā/ 'thus'; /kamō³/ 'we, inclusive of you' with /bunō³/ 'earlier in the day'; and /rumōh/ rhymes with /mēh-mōh/ 'confusion'. On the other hand /nūwɔ³/ 'you return' rhymes with /taŋɔə/ 'up to'; /lapa/ 'quickly' with /haba/ 'talk, news'; and /mon/ 'dew' with /putron/ 'bring down'. There are some-

times exceptions where a nasal vowel is made to rhyme with an oral vowel, but these comprise only a minority of cases: they do not appear to be any more numerous than other inaccuracies in the rhyme. So the intuitions of the poet as to what is a good rhyme correspond to the phonemic analysis of nasal vowels as distinct phonemes. This contrasts with orthographical intuitions and ‘minimal pair’ intuitions concerning the location of the contrast of NV with NV. If the poetic intuitions matched orthographical intuitions then these two sequences would be made to rhyme, because in the orthography they differ in their initial stop, which is not part of the syllable rhyme. However they do not rhyme, rather NV rhymes with CV and NV with CV. It is striking that although rhyme matches nasal vowels with nasal vowels, irrespective of the preceding consonant, orthographic intuition obscures the phonemic status of this nasality.

Discussion

Two characteristics of the orthographies need accounting for in that neither is reflected in a phonemic analysis based solely on the distribution of phones.

(1) There is conditioned markedness. In the orthographies, when the contrast in nasality is reflected in the spelling, the sequences CV and NV are unmarked, CV and NV are marked. After an oral consonant the *presence* of a nasal vowel is marked, after a nasal consonant the *absence* of nasalisation is marked. By markedness I mean literally orthographically marked by a more complex representation: CV and NV are always represented as a simple consonant + vowel sequence, with no further orthographic mark. However NV is always marked by including a voiced oral stop letter before the vowel, except in Abdul Gani Asyik’s analysis, where he underlines the nasal stop. CV, when it is distinguished from CV, is the marked member of this pair. In the traditional orthography only 2V is distinguished, by using the consonant letter ‘ain instead of *alif* or *hamza* for /2/. In Aceh, modern publications transfer this system to a Roman orthography by using an apostrophe for ‘ain and nothing for *alif* and *hamza*. Snouck Hurgronje’s orthography marks the CV sequence by using bold or italics for the consonant (except that 2V is represented as ’V and 2V as ‘V). Budiman Sulaiman’s orthography marks CV by an apostrophe before the vowel letter, to distinguish it as nasal. Abdul Gani Asyik uses a comma typed under the vowel for the same effect.

This patterning of orthographic markedness corresponds to the actual phonological markedness. In unstressed syllables only the unmarked sequences occur, thus never C \bar{V} or NV. Moreover the frequency of occurrences of the four sequence types in stressed syllables reflects the markedness: CV and N \bar{V} sequences are much more common than C \bar{V} and NV sequences, the rarest being the C \bar{V} type, and the commonest CV. C \bar{V} sequences are also particularly restricted in that the C is voiced only in expressives: /bē?/ 'goat's cry'. This distribution of frequencies of occurrence means that the orthographies are efficient in that they only give a special marking to the less frequent sequences, and it is redundant to mark vowel nasality (or its absence) in unstressed syllables. The markedness is also reflected in spelling errors or inconsistencies; when a writer is inconsistent in representing nasality the inconsistency always involves leaving out the mark. Thus Budiman Sulaiman (1977) often writes *han* instead of *h'an* for /hān/ and in Hasjim (1960) the sequence /No/ is often spelled (inconsistently) as *Nō* instead of *NCō*, thus *teumōn* 'fat' instead of *teumbōn* for /tumon/ and *pungōng* 'anus' (with a digraph *ng* for /ŋ/) instead of *punggōng* for /puŋŋōŋ/. Note that *ō* (for /o/) must be oral. The markedness is also manifested in the diachronic tendency for sporadic changes of NV > N \bar{V} , thus /pajāŋ/ 'tall' is attested in Acehnese instead of the expected /pajanŋ/ (cf. Malay *panjanj*). Phonetically C \bar{V} and NV are marked since they are characterised by a movement of the velum at the transition from consonant to vowel: the velum opens for C \bar{V} and closes for NV. In this respect these sequences are more complex phonetically than CV and N \bar{V} . It is of interest to consider how this markedness could be captured in a formal description of Acehnese phonology. The most obvious solution would be to represent vowels as [± nasal] in the lexicon only in the cases which are orthographically marked, i.e. for NV and C \bar{V} sequences. The vowel in sequences of the type N \bar{V} and CV would have a blank entry, literally be unmarked, for the feature [± nasal]. Nasality in these cases would be predicted by two markedness rules. The first given below accounts for the sequence CV, and is a general rule, valid for all languages:

$$(1) \quad U \text{ nasal} \longrightarrow -\text{nasal} / \left[\begin{array}{c} \text{—} \\ \text{v} \end{array} \right]$$

In other words, oral vowels are in general unmarked. The second rule is specific to Acehnese and other Indonesian languages (see Court 1970). It accounts for the sequence type N \bar{V} :

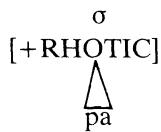
$$(2) \quad U \text{ nasal} \rightarrow + \text{ nasal} / N [\underline{V}]$$

In other words, after a nasal stop nasal vowels are unmarked. Constraints on the application of such rules would make these rules work to give the desired result: for example Kiparsky's Elsewhere Condition prevents these results from applying to the cases which are already specified in the lexicon for vowel nasality (Kiparsky 1982). The use of a conditioned markedness rule such as (2) which supplements a general rule like (1) is not without parallel. Kooij (1986) argues that such a rule is necessary to account for the markedness of [s] and [z] in Dutch. A general rule, which implies that in the unmarked case fricatives are – voiced, needs to be supplemented by a language specific rule which makes the value + voiced unmarked between vowels. It is a point in favour of the analysis of Acehnese markedness presented here that it provides a neat account of the relative frequencies of all four sequence types. As was noted above, the commonest is CV, then NV, then NV, and the rarest is CV. The sequence to which rule (1), the general markedness rules applies, is thus the commonest; the sequence accounted for by (2), the conditioned markedness rule, is the next commonest; then comes NV, the exception to the conditioned rule; and least common of all is CV, the exception to the general markedness rule. The general principle here is that the least marked case conforms to the most general rule, the most marked case is the exception to the most general rule. One would predict that if a language were to have only three of the four sequence types, it would be CV that is not represented, and this is indeed so (see Durie 1985:26). Also in Dutch, as described by Kooij (1986), there are exceptions to the conditioned markedness rule for fricative voicing, but not to the general markedness rule (other than those accounted for by the conditioned rule).

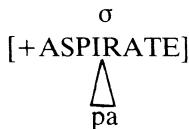
(2) A striking characteristic of the orthography of Budiman Sulaiman and the phonemicisation and orthography of Abdul Gani Asyik is the inconsistent location of the contrast. For CV: CV it is a vowel contrast, but for NV: NV it is a consonant contrast. The evidence of informants' intuitions accords with these analyses. The traditional Arabic-derived script and Snouck Hurgronje's orthography also agree in marking NV: NV as a consonant contrast, but the former only distinguishes a CV sequence in the case of ?V, and Snouck Hurgronje marked CV: CV as a consonant contrast in all cases, an approach which conflicts with native speaker intuitions and has not been followed by Acehnese who have described their own language.

The variable location of the contrast in the two native speaker analyses can be made sense of if nasalisation is represented in speakers' internal phonology in a way that is not strictly segmental. It is not implausible that it is represented *prosodically*, that is, as a feature of a whole syllable, or at least of the initial consonant + vowel part of the syllable. As such it would not be located on any one segment. Its realisation phonetically on the vowel can be understood as due to certain rules or conventions, not an inherent part of the representation itself.

There is no need to represent all phonological contrasts prosodically (i.e. *supra-segmentally*); one can impose a feature such as nasality (or its absence) over a segmental syllable structure. However vowel nasalisation is only one of many aspects of traditionally phonemic representation which lend themselves to a prosodical representation. Consonant clusters and aspiration in Acehnese could be treated prosodically, thus /pra/ might be represented (impressionistically) as:



and /p^ha/ as:



where RHOtic and ASPIRATE represent the prosodies of /r/ and aspiration respectively. It is not necessary for a phonological characteristic to be prosodic in the phonetic sense of extending over more than one segment in order for it to be represented as a prosody. In the process of speech formation the prosodic representation is linearised and realised sequentially on segments, and whether a component of the realisation is spread over one or more than one segment is determined by the rules or conventions which interpret the representation, not by the representation itself. In the case of the nasal prosody the realisation is primarily on the vowel (but also partially on adjacent consonant segments, see Durie 1985:15 [vi], 20 [iii]) and this constitutes the primary auditory mark. A phonemic analysis, based as it is upon the distribution of phonetic segments, must treat Acehnese as having contrasting sets of nasal and oral vowels.

Thus one can speak of two levels of representation: a hypothetical deeper one where nasalisation is not a segmental feature, and one much closer to actual articulation where the contrast is localised on the vowel. Rhyme is sensitive to this latter level, because it relies on how words sound. The discrepancy between the orthographic representations discussed here and the natural phonemic one can be understood as due to differing ways in which phonological representations can be accessed. Conscious orthographical intuition apparently accesses some features of the inner representation more directly, and does not just reflect the system of phonetic contrasts. If one assumes that orthographic intuition accesses an inner phonological representation at the point where nasality is prosodic, then it is understandable that the orthographic segmentation does not correspond to the actual phonetic segments, as in the orthographies of Budiman Sulaiman and Abdul Gani Asyik, for which the nasalisation contrast is located for some sequences on the consonant and for others on the vowel. In the traditional Arabic-derived orthography the contrast is mostly located on the preceding consonant, but sometimes not marked at all. The fact that it is not marked is not surprising if one assumes an underlying prosodic representation; instead of forcing an unlocated prosodic contrast onto a particular segment the orthography omits it altogether.

An intriguing question which remains is why the contrast is marked where it is in the orthographies: if an orthography must locate the contrast on a segment, why does it assign it to the consonant for NV and to the vowel for C^NV. In the case of the Arabic-derived script, influenced as it is by Malay conventions, the analogy of numerous Malay cognates is perhaps sufficient an explanation: A sequence such as /ma/ in Acehnese often corresponds to /mba/ in Malay. Moreover, the use of '*ain*' for /ə/ in sequences of the type ?V could be accounted for by the tendency in borrowing from Arabic to reflect '*ain*' with a following nasal vowel, as in Acehnese /do?ā/ 'to pray'. These reasons are not sufficient to explain Budiman Sulaiman's and Abdul Gani Asyik's analyses, nor the native speaker intuitions that I observed when eliciting minimal pairs. I can think of three possible reasons why the contrast should be located where it is:

(i) Phonetic Intuition. Because the phonetic contrast involves nasalinity the contrast is felt to reside on the nasal segment: the nasal stop in NV and the nasal vowel in C^NV. This explanation relies on the speaker having some access to his phonetic intuitions whilst forming his orthographic intuitions, although the phonetic awareness is dominated or filtered by the pro-

sodic nature of the inner representation.

(ii) Orthographic Efficiency. Perhaps a more plausible reason is suggested by a consideration of the alternatives: marking NV as having a marked type of vowel and C^V as having a marked type of consonant. The total number of syllable initial oral consonant segments is 26, with several consonant clusters possible; the number of oral vowels is 15; nasal vowels 12; and the number of nasal stops is 4. There are 4×15 possible NV sequences, so it is clearly more efficient to mark the four nasals than the 15 vowels as different. Similarly there are 26×12 possible C^V sequences, so again it is much more efficient to mark the 12 vowels as nasal than the 26 consonants, not to mention the consonant clusters. Snouck Hurgronje chose to mark the consonants as contrastingly nasal, with the practical consequence that many writers have found it convenient not to distinguish these ‘nasal’ consonants at all (e.g. Drewes 1979, Kreemer 1931) when using his orthography, or one derived from it.

(iii) The Markedness Rules.² Another possible reason for the orthographic location of the contrast is suggested by markedness rule (2), which specifies that after nasal stops nasal vowels are unmarked. Since the conditioning environment for this rule is the preceding nasal stop, exceptions to the rule are perhaps interpreted by a native speaker in terms of this preceding stop. If it can somehow be thought of as not fitting the specifications of the rule, then the rule can in some sense be saved from exceptions. This is perhaps what Abdul Gani Asyik is feeling for when he calls these ‘incomplete’ nasals. They are not ‘proper’ nasals because they do not condition nasality in the following vowel. Baihaqi et al (1977) describe a similar nasal series for Gayonese which they call ‘non-nasal’ nasals (*nasal taksen-gau*) in contrast to ‘nasal’ nasals (*nasal sengau*). What the native speaker’s intuition does is to save the markedness rule from exceptions. The many orthographies which transcribe an NV sequence by inserting a voiced oral stop before the vowel are also saving the rule; since in the orthography the vowel is not preceded by a nasal stop, it would follow from the markedness rules that it is not nasal. So much for the NV sequences. Abdul Gani Asyik and Budiman Sulaiman analyse C^V sequences as having contrastively nasal vowels. This does not need any special accounting for: it is after all a vowel contrast. The markedness rule to which these sequences are exceptions — rule (1) — has no conditioning environment, so there is no reason to look to the preceding consonant for the location of the contrast. Snouck Hurgronje’s orthography, however, does attribute the contrast to the preceding

consonant by distinguishing a set of nasalised oral consonants. This is taking the rule-saving strategy to its extreme limits. In his orthography the markedness rules admit of no exceptions since for C^NV the preceding consonant is treated as nasal, and for NV the vowel is preceded by a voiced oral stop letter.

The important general point of this study is that a segmentation based on conscious intuitions, as in orthographies of Acehnese, where writers exercise a considerable amount of individual choice, may be at variance with the segmentation implicit in the phonetics, because conscious accessing of inner phonological representation can be dominant for these intuitions. This is a well known phenomenon for languages with complex morpho-phonemics, where underlying segmental representation is quite at variance from its surface manifestation, but Acehnese presents a quite different case, where the inner representation is apparently not segmental in the first place, but prosodic, and it reflects a markedness which is obscured by the phonemic, segmental analysis. It is not a question of the orthography reflecting an underlying segmentation, but of juggling a prosodically coded characteristic, along with its markedness, into the orthography.

Unordered Segmentation - Evidence from English and Georgian

It was suggested above that nasalisation in Acehnese is not located on any particular segment in the phonology, but rather on a larger unit such as the syllable, and that even whole segments such as /r/ in a consonant sequence could be analysed in this way. Instead of being regarded as a phoneme in sequence with other phonemes, the /r/ would be analysed as something like syllable 'rhoticisation'. The realisation of the segment in sequence would not be an inherent part of the lexical representation. There are two quite different pieces of evidence from English and Georgian to support the hypothesis that speakers do actually represent some segments without ordering them in the sequence in which they appear in speech. Gil and Radzinski (1986) point out that Georgian has considerably free variation in consonant ordering. The following are all acceptable variants of the same lexical item, glossed as 'trainer':

mc'vrtneli	mvc'trneli	c'mrtvneli
mc'rtneli	mrc'vtneli	c'rmtvneli
mc'trvneli	mrc'tvneli	mc'rtnveli
mvc'rtneli	rmc'vtneli	mrc'vnteli

This would appear to be strong evidence for unordered syllable structure. The freedom of consonant ordering would then not be a characteristic of the lexical representation, which would have unordered segmentation, but rather of the realisation rules by which this representation is transformed into a linear array. All these variants would have a single (?) lexical representation, which is as it should be, since speakers of Georgian recognise them as the same word.

The example from English involves a type of spelling error described by Marcel (1980) for English. He shows that some people, English speakers, have trouble spelling certain consonant sequences. For initial clusters a liquid was inconsistently omitted or misplaced, and the voicing of the stop sometimes reversed. Sometimes with a cluster like /skr/ the liquid and the /s/ were omitted and the stop voiced, thus *scrape* spelled *gape*. Examples of a misplaced or omitted liquid are *glass* spelled *gasse*, *blood* spelled *boold*, and *grass* spelled *cass*. For syllable final sequences of sonorant + stop the order of the consonants was sometimes reversed, or one of the consonants was sometimes left out. Thus *bent* as *betn*, or *bet*, and *hummed* as *hum*. Marcel suggests that the errors are because the people with this problem are using a different set of features from normal adults to analyse and produce speech. In particular, he suggests that “normal adults, for whatever reason, classify in terms of phonemes while the deviant spellers classify in terms of distinctive features, reducing syllables to CV- and -VC with concatenated additions” (p. 398). This is more or less how a prosodic phonology can be structured. Upon a basic syllable structure, elements such as sonorants in consonant sequences and consonantal voicing can be encoded prosodically. In the representation of words such segments as /s/ in *scrape* and /r/ in *grass* would not necessarily be located in sequence as in the phonemic representation, but rather as features of the syllable. If the problem spellers are reflecting such a representation then their errors of wrong orderings and omissions are immediately understandable, since the representation would not have the misspelled elements located in their phonetic sequence; they would be prosodic characteristics and thus unordered in the representation. It is not necessary to assume that such problem spellers have a different set of phonological features, as Marcel suggests; it is possible that only the way in which they access their phonologies differs from ordinary adults. It is quite plausible that normal adults have both an inner prosodic phonological representation and another level where their prosodies are realised segmentally. What distinguishes the normal adult from

the problem speller is that the latter experiences difficulty accessing the fully ordered realisation, and so his spelling reflects the incompletely sequential prosodic representation. The problem appears only to be manifested in spelling: it is significant that Marcel could observe no differences in the speech of the problem spellers. This is quite consistent with the proposition that they have a normal phonological representation but problems in consciously accessing the order of segments for the purposes of writing. A difference in the Acehnese spelling ‘problem’ is that it is apparently normal to have difficulty in accessing the phonetic location of vowel nasality when spelling. The prosodic representation of nasality in Acehnese is more dominant than that affecting English consonant sequences. It is striking that for Acehnese, rhyme patterns according to the phonemic analysis of nasality, so the difficulty in segmenting nasality for orthographical representation is quite a separate matter from the ability to determine whether two vowels are the same for the purposes of rhyme. It is to be expected that the English speakers with the spelling problem that Marcel describes would also have no more difficulty telling whether *bent* and *bet* rhyme, even though they might have trouble distinguishing them in spelling.

NOTES

1. All Acehnese forms given here are from the dialect of Acehnese spoken in Aceh Utara near the town of Bireuen. I conducted field work there over several months in 1981, 1982 and 1983. Acehnese terms (not example words) given in italics are in the orthography used in Durie 1985.
2. This explanation was suggested to me by Thore Pettersson.

BIBLIOGRAPHY

- Abdul Gani Asyik (spelled *Asjik* on the title page). 1972. *Atjehnese morphology*. Master's thesis, Institut Keguruan dan Ilmu Pendidikan Malang.
- Baihaqi A.K. et al. 1977. *Struktur Bahasa Gayo*. Banda Aceh: IAIN Jami'ah Ar-Raniri and Universitas Syiah Kuala.
- Budiman Sulaiman. 1977. *Bahasa Aceh*. Vol. 1. Bireuen, Aceh Utara, Indonesia: Mahmudiyah.
- Court, C. 1970. “Nasal harmony and some Indonesian sound laws.” In:

- Pacific Linguistics Vol. 13: Pacific Linguistic Studies in Honour of Arthur Capell.* Linguistic Circle of Canberra.
- Drewes, G.W.J. 1980. *Two Achehnese poems.* Bibliotheca Indonesica 20. The Hague: Martinus Nijhoff.
- Durie, M. 1985. *A grammar of Acehnese on the basis of a dialect of North Aceh.* Dordrecht: Foris.
- Gil, D. and D. Radzinski. 1986. "Georgian syllable onsets: some arguments for unordered hierarchic phonological structures." In: *Phonologica 1984. Proceedings of the Fifth International Phonology Meeting, Eisenstadt, 25-28 June 1984.* Cambridge: Cambridge University Press.
- Hasjim, M.K. 1969. *Himpônan hadih madja.* Banda Atjeh: Dinas Pendidikan Dasar dan Kebudajaan.
- Kiparsky, P. 1982. "From cyclic phonology to lexical phonology." In: H. van der Hulst and N. Smith (Eds.) *The structure of phonological representations I.* Dordrecht: Foris.
- Kooij, J.G. 1986. "Phonological markedness and morphological alternations." In: *Phonologica 1984. Proceedings of the Fifth International Phonology Meeting, Eisenstadt, 25-28 June 1984.* Cambridge: Cambridge University Press.
- Kreemer, J. 1931. *Atjëhsch handwoordenboek.* Leiden: E.J. Brill.
- Langen, K.F.H. van. 1889. *Handleiding voor de beoefening der Atjehsche taal.* 's Gravenhage: Martinus Nijhoff.
- Marcel, T. 1980. "Phonological awareness and phonological representation: investigation of a specific spelling problem." In: U. Frith (Ed.) *Cognitive processes in spelling.* London: Academic Press.
- Snouck Hurgronje, C. 1900. "Atjëhsche taalstudiën," *Tijdschrift voor Indische Taal-, Land- en Volkenkunde* 42, pp. 346-442.

PHONOLOGICAL RECODING IN THE READING PROCESS

Hartmut Günther

1. *The problem*

OMNE VERBUM SONAT — every word sounds.¹ This dictum is found in Augustinus' work 'De Dialectica'. Reading the passage where it occurs more closely quickly reveals that it is based on a still earlier work, namely, Aristotle's essay 'Peri Hermeineias', from which the following quotation is drawn:

ἔστι μὲν οὖν τὰ ἐν τῇ φωνῇ τῶν ἐν τῇ ψυχῇ παθημάτων σύμβολα, καὶ τὰ γραφόμενα τῶν ἐν τῇ φωνῇ.

Gelb 1963, in his seminal book on the history of writing, gives the following translation (p. 13):

Spoken words are the symbols of mental experience and written words are the symbols of spoken words.

As so often with translations from ancient Greek, this is rather an interpretation. However, since it clearly indicates how Aristotle has been read by most people (including, for instance, Augustinus), the following model (1) might be called 'Aristotelian' despite the fact that presumably Aristotle's view of writing was different (see Günther 1983a:25f., Haas 1970:10f. for discussion):

(1) written symbol → spoken symbol → meaning

Model (1) represents a theory of written language which holds that the written symbol represents the sound symbol, and the sound symbol represents the meaning. This model has been adopted by most scholars writing about written language, and it is also the model the man in the street has in mind when speaking about writing. In the linguistic sciences, one may well maintain that this model prevented further discussion about written language since it seems so obvious that further investigations appeared to be point-

less (see Coulmas 1981, reviewed in Günther 1982, for a general discussion). I have argued in several articles that model (1) is not an adequate account of the structure of written language (Günther 1981a,b, 1983a, 1984). In the present paper I wish to discuss the fact that a psychological interpretation of (1) is also faulty. According to proponents of model (1) in reading research, there is only one way of reading (see Gough 1972 for a very decisive statement of this sort): The written symbol is perceived, internally transformed into a sound symbol (usually a chain of phonemes), and the sound symbol is used in order to access meaning. Only if one knows how a sequence of letters (actually, also a logograph!!) sounds is one able to find out what it means. OMNE VERBUM SONAT — a word is a word only in sound.

2. *Three schematic models of the reading process*

The question of whether a structural model like (1) can also be given a psychological interpretation has been discussed since the beginning of experimental reading research at the end of the last century (see the classical monograph by Huey 1908). Most experiments performed with this goal in mind follow the same general scheme. It is assumed that if phonological effects show up in purely visual reading tasks, then this would be evidence in favour of a model like (1). Perhaps the oldest observation counted as evidence of this sort is subvocalising, which most readers do quite automatically. Still, there are two important objections against such a view. First, it is rather easy to learn to avoid subvocalising by certain feedback techniques (see Hardyck & Petrini 1970 for example). Second, it is not clear if subvocalising does in fact occur before lexical access, as displayed in (1). Compare the following three schematic reading models:

- (1) written symbol → sound symbol → meaning
- (2) written symbol → meaning → sound symbol
- (3) written symbol ← meaning ← sound symbol

In order to prove that model (1) is correct, the mere demonstration of phonological effects in purely visual tasks is insufficient since models (2) and (3) could account for such effects as well. They are, however, not included here for the sake of the argument only, but rather since many scholars in reading research believe them to be correct instead of model (1).²

The virtue of modelling the reading process schematically in this way is that one can design experiments whose outcome is incompatible with some of the models, thereby allowing an empirically founded decision among them. Before turning to the available evidence, I wish first to be a bit more precise about the concept of 'phonological recoding' and to introduce the concept of the 'mental' or 'internal lexicon'.

3. Phonological recoding and the internal lexicon

Generally speaking, phonological recoding in reading is the process by which written symbols are transformed into sound symbols. In alphabetic writing, the linguistic description of this process is achieved by a set of grapheme-phoneme-correspondence rules (GPCs, see Venezky 1970 and Bierwisch 1972 for some descriptive approaches to these rules in English and German). Since, however, linguistics deals with matters of the structure of language, whereas reading is a psychological process, to be dealt with by psychology, it is inadmissible simply to assume that linguistic statements can be extended to the psychological domain without justification (see Henderson 1982: part I for a brilliant and devastating critique of this approach). Hence, it seems appropriate to reformulate the definition of phonological recoding: Phonological recoding is the process by which internal representations of written symbols are transformed into internal representations of sound symbols.

Support for the notion that phonological recoding occurs in reading comes mainly from three not completely independent arguments. First, it is assumed that linguistic rules can be interpreted as psychological processes. Henderson's critique of this approach has already been mentioned. Second, in alphabetic writing systems, the skilled reader is able to read words aloud he does not know. I shall return to this point later. Third, by some kind of simplicity argument, one might hold that in listening speech is internally represented as a chain of phonemes. Since rules can be written which transform graphemes into phonemes, the only achievement necessary in order to learn to read would be to learn these rules, since phonemic representations are necessary anyway for speech perception. The trouble with this argument is that the basic assumption is wrong. It is by no means clear nowadays how speech perception is in fact achieved. But that it does not work simply by a transformation of the speech signal into a chain of phonemes seems evident. In fact, the most prominent attempt (Liberman, Cooper,

Shankweiler & Studdert-Kennedy 1967) to prove such a theory of speech perception as a sequence of different stages, whereby the acoustic signal is finally transformed into a chain of phonemes, led to precisely the opposite conclusion.

At present, it is unfortunately rather unclear what the nature of internal representations of speech events really is. Therefore, I do not wish to make any specific assumptions about their nature. Rather, I prefer the following convention (following Tillmann 1980, see also Messing, Günther & Kegel 1980): phonemes are not mental entities which underly phonetic realizations; rather, phonemes are used to represent phonetic events. By phonological recoding in reading, then, I would like to denote a process by which internal representations of what is seen are transformed into internal representations of phonetic events, which can be represented by chains of phonemes.³

The main problem of reading research directed towards the question of phonological recoding can be found in the distinction between direct and indirect access to the mental lexicon, which has been elaborated by Coltheart and his coworkers. In Coltheart, Davelaar, Jonasson & Besner 1977, the following definition is given:

This lexicon consists of all the information a reader or listener has acquired about the words of his language — in particular information about the meanings, the pronunciations and the spellings of the words he knows. The lexicon is made up of subsystems of information, each subsystem characterizing a particular word. These subsystems are lexical entries.

...

Within this framework, reading a word for meaning consists of extracting some form of information from a set of ink marks and using this information to reach the word's lexical entry. The successful execution of this process ("lexical access") furnishes the reader with the meaning of the word. (536)

There are several problems with the concept of the mental or internal lexicon (see Hörmann 1983 for a cogent critical discussion). Also, it is rather likely that the concept of one lexical entry for all kinds of information (phonological, semantic, syntactic, etymological, frequency, register, etc.) is misleading. Still, for the purposes of the present paper, the concept of the lexical entry proposed by Coltheart et al. 1977 is sufficient. The problem of phonological recoding just defined can now be rephrased like this: Is lexical access in reading direct or indirect? That is, can the lexical entry be accessed directly by recognizing a written symbol, or is it always (or sometimes) necessary that the written symbol be phonologically recoded before

the lexical entry can be reached? It becomes clear now that the distinction between direct and indirect access is concerned with the locus of the phonological recoding operation. As models (2) and (3) show, phonological recoding may also occur after lexical access; hence, direct access and phonological recoding are not mutually exclusive. Therefore, some authors use the expression "prelexical phonological recoding" to refer to indirect lexical access.

4. The discussion of indirect access to the mental lexicon in reading research

Experimental reading research concerning the question of phonological recoding concentrated from the beginning (see Huey 1908 for a review) on the recognition of single words. For the technical and methodological reasons see Günther 1983b: Chpt. 1; it might suffice to say here that single words are easier to present visually and that experiments can be better controlled when words are presented instead of sentences.

A first landmark still worth considering was set by Eleanor Gibson and her coworkers. Gibson, Pick, Osser & Hammond 1962 were the first to show by observing modern experimental standards that in tachistoscopic word recognition experiments pseudowords are better recognized than non-words. The former are pronounceable letter strings, the latter are illegal strings, anagrams of the former (cf. the pair GLURCK - CKURGL). Gibson et al. concluded from this result (which has been replicated over and over again) that in visual word recognition, the central point is context-sensitive GPCs.

... The proper unit of the reading process is neither the single letter nor the whole word but a higher order invariant derived from grapheme-phoneme-correspondences. Rules can be drawn up for predicting spelling from sound if temporal patterning of 'clusters' is taken into account. It was proposed that the skilled reader has learned to perceive as units the letter patterns conforming to these rules and that such patterns have, therefore, an advantage in reading. (p. 570)

Three points against the conclusion that this result bears evidence for model (1) can be made. First, as observed by Gibson, Shurcliff & Yonas (1970), the effect of pseudoword superiority is also found in congenitally deaf subjects, where the assumption of phonological recoding is rather unlikely. Second, as observed by these authors as well as by Baron & Thurston 1973, the effect might entirely be due to orthographic (i.e. visual)

regularity of the pseudowords against non-words.⁴ Third, it might be that these clusters are learned because they are based on the phonology of the word, but that they are used by the skilled reader without any reference to phonology. This seems to be a conclusion reached by many modern authors.

Another very influential attempt to support model (1) by empirical evidence is found in several articles by E.E. Smith and K. Spoehr. They demonstrated that the recognition of words and pseudowords in tachistoscopic experiments depends on the syllabic structure of the stimuli. The most impressive evidence is presented in Spoehr 1978. Here, she found a solid effect of syllabic and of phonemic length on visual word recognition. In particular, five letter words were better recognized when the five letters represented only 4 phonemes (SHARK) instead of 5 (START). The logic of her conclusion seems straightforward: Phonological recoding, occurring mandatorily, leads to fewer errors in 4 phoneme words because of the shorter phonemic structure.⁵

There are many experimental reports of this sort, most of them briefly summarized in McCusker et al. 1981 and other reviews. The trouble with these is the fact that whereas they clearly show some influence of phonology in visual word recognition, they are not sufficient to prove model (1), i.e. indirect lexical access. The point has been stated most clearly in Coltheart 1978: if one is to investigate how lexical access occurs, one has to make sure that it does in fact occur. Neither in Gibson et al. 1962 nor in Spoehr 1978 is it guaranteed that lexical access occurred during the experiments, since saying what letters have been perceived after a short tachistoscopic presentation does not require lexical access. Even the word superiority effect and the effect found on words in Spoehr 1978 is no proof: It might well be that the phoneme length effect only occurs when subjects treat words not as words but as mere letter strings.⁶

5. *Lexical decision experiments*

An experimental task which necessarily implies lexical access is the lexical decision task, first used by Herbert Rubenstein and his coworkers at the beginning of the seventies. In this task, subjects are to decide as quickly as possible if the visually presented letter string is a word. The dependent variable is the reaction time (RT), sometimes also the number of errors. If the stimuli for the NO-answer are pseudowords, i.e. letter strings which are

orthographically legal and which are pronounceable in the language of the subjects, lexical access in the experiment necessarily occurs. There is only one way to find out if e.g. SLINT is an English word or not, namely, to search the internal lexicon for such a word. If, then, phonological effects in this paradigm can be demonstrated, they seem to imply that prelexical phonological recoding occurs.

Rubenstein, Lewis & Rubenstein 1971 were the first to claim to have achieved such a result. They presented pseudowords like SLINT and pseudohomophones like BRANE, which are also not English words, but which are pronounced like English words (BRAIN). They found that lexical decisions for pseudohomophones (BRANE) took significantly longer than for pseudowords of the SLINT-type. In terms of model (1), the explanation of the pseudohomophone effect is straightforward: since all letter sequences are phonologically recoded before lexical access, BRANE first appears to be a word. Only by a spelling recheck does the subject find out that it is nevertheless not an English word.⁷

The pseudohomophone effect has been replicated several times since Rubenstein et al.'s 1971 report. However, the argument that it provides evidence in favour of model (1) has been questioned also. In particular, McQuade 1981 and Gfroerer, Günther & Weiss 1984 were able to demonstrate how to get rid of the effect in a lexical decision task: If most stimuli for the NO answer are pseudohomophones, the effect disappears. It seems that subjects in this condition somehow realize that the phonological code is unreliable in this situation and simply ignore it. In Gfroerer et al. 1984, further evidence for this point of view is adduced by the observation that in the condition where predominantly pseudohomophones are present for the NO-answer, the RT for the YES-answer is quicker than in the condition where only few pseudohomophones occur. The conclusion seems to be warranted that subjects use different strategies in these conditions, relying, as it were, to a greater or lesser extent on phonological codes. This conclusion, however, is definitely incompatible with model (1).

Even more damaging for the position that the pseudohomophone effect proves model (1) is an argument brought forward by Coltheart 1978. The problem is that the pseudohomophone effect is achieved in the NO-answer of a lexical decision task. Hence, it is not clear if the effect can be generalized to words. Consider the following account of the pseudohomophone effect. It might be the case that phonological recoding only occurs when the visual search, i.e. the direct access, fails to be positive, and only in this case do subjects turn to a phonological recoding strat-

egy they never use with words. The well attested finding that pseudowords yield longer latencies than words is a case in point, see also Günther 1983b for an analysis of latencies for several types of nonwords. Therefore, Coltheart 1978 argued, what is needed is positive evidence, i.e. the demonstration of phonological recoding in words, in the YES-answer of lexical decisions.

One influential attempt to find evidence of this sort stems again from Rubenstein et al. 1971. They constructed a rather ingenious argument about why the lexical decision for the less frequent member of a homophone-pair like WEEK/WEAK should show a longer RT than an appropriate non-homophonic control word. I quote from Coltheart et al. 1977, who echo Rubenstein's original argument:

In the case of words, a string will only suffer through being a homophone when it is the *less* common member of a homophone pair. Consider the words WEEK (frequency 1278) and WEAK (frequency 243). Whichever of these two is actually presented, the lexical search will (in the case of phonological recoding, HG) first stop at WEEK, the more frequent word. If indeed WEEK was presented, the spelling check will succeed and the response "yes" will be made; the existence of another lexical entry with the same phonological code further down in the lexicon, will have no effect. But if WEAK is presented, there will have been a time-consuming stop, spelling check and restart during the search down to the entry WEAK. So WEAK will suffer by being a homophone, while WEEK will not. (S.541)

The evidence available is ambiguous. Coltheart et al. 1977 failed to achieve the predicted effect which is, on the other hand, present in Rubenstein et al. 1971 and in Davelaar, Coltheart, Besner & Jonasson 1978. The latter authors, however, were also able to show that the homophone effect disappears if one employs the method to destroy the pseudohomophone effect discussed above: If very many homophones are presented in a lexical decision experiment, no homophone effect is achieved. Humphreys, Evett & Taylor 1982, in a different paradigm, found that error scores are significantly lower in tachistoscopic word recognition when one member of a homophone pair is preceded subliminally by its counterpart, i.e. if WEAK is preceded by WEEK, displayed so briefly that subjects are not able to report anything of the subliminally presented word. Underwood & Twaites 1982, on the other hand (paradoxically in the same issue of *Memory & cognition*), found an inhibition effect by presenting a wrong homophone member subliminally in peripheral vision when the target word was centrally presented. That is, they presented a word like KING in central vision

and, peripherally, they presented the wrong homophone member of a phonologically related word, i.e., RAIN instead of REIGN, for a short period of time. Despite several problems in their design, this result is damaging for model (1) since if lexical access relies primarily on a phonological code, subliminal priming should be possible even with the wrong homophone member resulting in facilitation rather than in inhibition for the target. In an unpublished experiment, using the same stimuli as Underwood & Twaites 1982, I replicated their main results in a double lexical decision task. It was more difficult to decide that KING/RAIN are both words than to say that BEAR/GREEN (where both meanings of the homophone are not related to GREEN) or THING/SMILE are both words. Summarizing the picture from experiments with homophones, one can say that only some authors produce evidence compatible with model (1), whereas others report definite counterevidence.

A similar situation is found in experiments where orthographically irregular words are used, i.e. words which apparently violate general GPCs. Consider the word GAUGE. It is unique in English in that the letter sequence AU must be translated into /eɪ/. If phonological recoding occurred prelexically, GAUGE would be translated into /gɔ:dʒ/. There is of course a lexical entry for this sequence of phonemes — but it is the wrong one as revealed by a spelling recheck (see fn. 5.). Search should therefore restart, no other entry would be found, and some kind of an exception dictionary need to be consulted. The available evidence is again ambiguous. Coltheart et al. 1977 and Coltheart, Davelaar, Besner & Jonasson 1979 did not achieve an irregularity effect. Bauer & Stanovich 1978 found the predicted effect in the normal lexical decision task; however, the effect vanished when subjects were set under time pressure.⁸ Parkin 1982 got an irregularity effect with “very irregular” words, but not with words which are only “moderately” irregular. The situation is further blurred by Glushko’s 1981 experiments where irregularity and consistency are pitted against each other. Hence, one may once again conclude that no evidence for model (1) is available, and that evidence for the indirect route of model (3) is not unambiguous.

6. *The pseudoword argument and data from neuropsychology*

A sceptic, quite willing to agree with the arguments hitherto provided, might ask now the crucial question: How, then, is it possible that the skilled

adult reader is able to pronounce words or letter strings he or she has never seen before? Our ability to pronounce pseudowords seems to imply that some kind of nonlexical⁹ GPC-procedure is necessary in order to arrive at the correct pronunciation. This argument is faulty in at least the following two points. First, reading aloud is a special form of reading, involving necessarily the transformation of print into sound. Although reading aloud is presumably historically older, it is nevertheless the less-used way of reading in our times. Furthermore, it is also ontogenetically not at all clear if learning to read necessarily involves as a first step the ability to transform written symbols into sound symbols. The evidence brought forward by Steinberg & Harper 1983, "Teaching written language as a first language to a deaf boy" (starting at the age of 2;5!!), clearly denies this necessity, see also K.B. Günther's 1982 review of the role of written language in the education of deaf children. The second objection against the unqualified pseudoword argument has been stated most clearly by Henderson 1982 and 1985: it is not at all clear at which stage in the process of transforming perceived print into articulated sound phonology comes in, and it is also unknown how it comes in. The general assumption that because alphabetic writing systems are alphabetic, pseudowords have to be pronounced via the application of GPC's, is not only methodologically invalid, but also shortsighted. Henderson (1982: part II) convincingly argues that pseudowords can be pronounced without any reliance on GPCs by use of lexical analogies. To give a rather oversimplified example: the pronunciation of a pseudoword like SLINT in English might be achieved by some kind of merging of the lexically (!) derived pronunciations of the words FLINT and SLINK. That is, because subjects know how similar words are pronounced, they are able to pronounce pseudowords. One should however mention the fact that whereas Henderson's argument is quite compelling, his only empirical evidence is found in the results of Glushko 1981. Since in 1985 Henderson himself mentions problems in replicating these findings, we are left with a brilliant argument without much empirical backing. Hence, further empirical evidence is urgently needed.¹⁰

The pseudoword argument also plays a central role in the last line of argument to which a proponent of the existence of nonlexical processes in phonological recoding might resort. The evidence to be discussed only briefly here stems from neuropsychology. Several case reports have been published describing alexic¹¹ patients who lack entirely the ability to read pseudowords aloud, whereas words can be read aloud without much diffi-

culty, and word comprehension is also not impaired. A very instructive overview of the main syndromes in alexia and their alleged implications for reading models is found in Patterson 1981. The interpretation of syndromes where the reading of pseudowords seems to be more or less impossible is straightforward for supporters of model (1) as well as model (3). In these patients, it is argued, the ability to use GPCs, i.e. the nonlexical route, is impaired, whereas the direct route is still intact. The existence of a syndrome which is in a sense opposite, namely letter-by-letter reading, seems to support this line of argument if one assumes that here the direct route is selectively impaired.¹² Still, there are at least two objections to this view. First, the last mentioned syndrome in particular points heavily to a possibility hardly ever mentioned in neuropsychological case reports, i.e. that of compensatory strategies. It might well be that letter-by-letter readers simply develop a strategy to overcome their reading disability which is definitely never used by adult skilled readers. In fact, I simply do not believe that the overt behaviour of pathological cases of cognitive systems so complicated as the human one appears to be can be used as a mirror of the normal functions since the ability to overcome impairments by availing oneself of different strategies or by acquiring new ones is in fact one of the main traits of the human cognitive system. Therefore, a direct jump from pathological cases to the normal by simple subtractive calculations seems to me completely unwarranted.¹³ The second objection is based on Henderson's rebuttal of the pseudoword argument. The conclusion that the inability to read pseudowords implies the impairment of a nonlexical GPC route to pronunciation is not valid. An alternative explanation would hold that these patients lack the ability to merge the pronunciation of words similar to the pseudoword presented in order to arrive at a correct pronunciation. It is perhaps useful to remark that I do not consider neuropsychological data as pointless. On the contrary, I share the present interest in this type of evidence. Still, methodological as well as theoretical flaws in the way arguments are evaluated in this research are also obvious. For impressive volumes of neuropsychological research on reading, the reader is referred to Coltheart, Patterson & Marshall 1980 on deep alexia and 1985 on surface alexia (see fn. 11).

7. *Interlude: Orthographic recoding in speech perception*

It is perhaps useful, before turning to conclusions, to break off for a moment. It was shown in the preceding sections that there are many experiments in visual word recognition whose outcome indicates that some phonological activity is going on. Nevertheless, evidence is scarce for the position that these effects occur prelexically as assumed mandatorily by model (1). It is surprising to see that the *experimentum crucis* for the existence of prelexical phonological recoding is still lacking despite the considerable efforts to achieve it during the last 15 years. This very fact, it seems to me, points to a general flaw in one of the main questions asked in visual word recognition research. As pointed out above, the logic behind most experiments was to demonstrate phonological effects in purely visual tasks. By the same logic, one could claim that orthographic effects in purely auditory tasks show orthographic recoding in speech perception. There are some reports of this sort available. Derwing & Dow (in this volume) found that phonological judgements are influenced by orthography. A very elegant experiment is reported by Seidenberg & Tanenhaus 1979. Subjects were asked to decide as quickly as possible which of three words rhymes with a previously presented word. There were two classes of targets: words which were orthographically similar and words which were not. For example, subjects were given first the word BOAT, followed by MOUSE — GOAT — FAIR in the similar condition, and by MOUSE — VOTE — FAIR in the dissimilar condition. The result of this purely auditory task — no writing was present, RT was measured by hitting a response key — was that the RT to classify VOTE-words as rhyming with BOAT was significantly longer than the RT to GOAT-words. The effect is a genuine one — it has been replicated by removing some possible artifacts by Nolan, Tanenhaus & Seidenberg 1981. If one were to interpret this result the same way as the proponents of the Aristotelian model interpreted their experiments, one would be inclined to argue that in speech perception, prelexical orthographic recoding occurs mandatorily, which is clearly absurd.

8. *Conclusions*

To summarize: Whereas the evidence definitely rules out model (1) as a model of the reading process, and evidence for the existence of a prelexical or nonlexical route from print to sound or to lexical access is scarce, abundant evidence is nevertheless available that phonological recoding

occurs quite regularly in visual word recognition. Whereas research seemed to concentrate nearly completely on the question of *when* phonological recoding occurs, surprisingly little attention has been given to the question of *what* the code looks like and *why* phonological recoding occurs at all. The confusion about the nature of the phonological code is most obvious when the number of labels is considered which are used for the concept I defined above and labeled "phonological recoding". McCusker et al. 1981:218, stating the same point, cite "speech recoding", "phonetic recoding", "phonemic recoding" and "deep phonemic recoding"; I may add "articulatory recoding", "auditory recoding" and "sound recoding" (see also fn. 2 above.). As these authors rightly observe, "opting for this term (phonological recoding-HG) begs rather than answers some fundamental questions, and a high priority should be assigned to studies aimed at more precisely describing the nature of the internal representation". The literature published thusfar, however, does not display many attempts of this kind.

Even less attention has been given to the question of why phonological recoding occurs at all, presumably because of the influence of model (1) in linguistic and psychological research. Nevertheless, it is somewhat annoying that the only evidence cited in this context is usually Conrad 1964, a rather simple experiment which studied auditory vs. visual confusions in a task involving the learning of letter names. This is a task rather far away from reading, and Conrad's result surely cannot support the conclusion usually cited, namely that written language is stored in short term memory in an articulatory code.¹⁴ Scheerer 1978 speculates that this may well be the case because a phonetic or phonological code might be more suitable for language than a visual code, but again, the adduced evidence is scarce.

Perhaps most surprising, however, is the fact that the very impressive literature on speech perception of the last 20 years is hardly ever mentioned at all in visual word recognition research — Henderson (1982: chpt. 2) is a notable exception. I strongly believe that a research program to resolve the phonological recoding issue should turn much more to the analysis of speech. The line of reasoning displayed by Derwing (in this volume) and Seidenberg & Tanenhaus 1979, who raise the question of how written language and orthography can influence speech production and speech perception, is well worth following. While speculating that their results are due to the metalinguistic nature of the tasks used, it would be of great interest to find out if influences of orthography can also be discovered in everyday speech. This is an area for experimental phonology, argued so convincingly

for by John Ohala in his paper at the Eisenstadt meeting. At least some first evidence is available. Morais, Cary, Alegria & Bertelson 1979, for instance, have shown that illiterate adults are virtually not able to perform the task of omitting or adding a single phoneme to words or pseudowords, a task which can be performed without problems by literate adults of the same population. Here it becomes obvious that the writing and reading skills influence the language abilities of the individual in general. Hence, interactions between reading and listening, speaking and writing are perhaps the most interesting fields of study. As far as the issue of phonological recoding is concerned, it seems inviting to speculate that the internal representation of speech changes when reading is learned. And it is not at all impossible that the phonological code used in reading is only used there. To resolve these problems, however, a drastic change of emphasis in reading research is needed.

A final remark. If linguistics (and phonology as a branch of it) is interested in describing the language system, its object, as an entity which is made by man, used by man, then it is necessary to acknowledge that modern man uses language also in its written form; and it is obvious that the use in its written form is not directly derived from speech. A literate adult in our modern societies is always not only a speaker/hearer, but also a writer/reader. The language system he uses comprises both modalities, the written as well as the spoken. OMNE VERBUM SONAT — no doubt about it, SED ETIAM SCRIPTUR....

NOTES

1. Eckart Scheerer drew my attention to this quotation. This review owes much more than can be written (!) to the several discussions we had on the issue of phonological recoding in reading.

The present paper is a thoroughly modified English version of my "Antrittsvorlesung", held during the process of Habilitation at the Philosophical Faculty II (FB 14) of the University of Munich (18.7.83).

2. Model (2) seems to be implied by Glushko 1981 and also advocated by Henderson 1982. Several versions of model (3) have been discussed in the literature under the heading of the term "dual access" models, esp. in neuropsychology — see McCusker, Hillinger & Bias 1981 for a review.

3. Note that this definition lacks much of the plausibility inherent in the informal definition at the beginning of this paragraph. Since the nature of internal representation of speech events is unclear, it is not useful to use any of the following terms for the process of phonological recoding: articulatory, auditory, phonemic, phonetic, since all of them imply specific assumptions

about the nature of these representations. I wish to stress that by using the term "phonological recoding", I do not wish to imply that these representations are phonemic.

4. It is of course true that orthographic regularity is at least partially dependent on phonology. The regularities of syllable initial or final letter clusters, in fact the definition of the syllable itself is most easily done in terms of concepts like "vowel letter", "stop letter" etc. (see Günther 1983a for discussion). Still, these are again structural and not psychological facts. That is, even if it turns out that linguistically the orthography of a language like English or German is best accounted for by a system of GPC rules dependent on the regularities of phonology (as proposed by Bierwisch 1972), this is not a valid proof of the perception of orthographic regularities via phonology by the skilled reader. Furthermore, as Weiss, Günther & Gfroerer 1984 were able to show, the mere regularity of stimuli is not the true reason of the word superiority effect as first claimed by Baron & Thruston 1973. Rather, it seems that lexicality and regularity are additive, as proved by showing that orthographically regular abbreviations like USA are recognized more easily than irregular abbreviations like IBM despite the fact that both types of abbreviations are not pronounced via GPCs, but letter by letter.

5. It is perhaps worth mentioning that Spoehr was also able to demonstrate how to get rid of the effect. Still, her particular explanation about why this should be so seems rather ad hoc.

6. See Coltheart 1978: 170f. for the elaborate statement of this critical argument. Furthermore, neither of the tasks mentioned by Coltheart (tachistoscopic word recognition, letter cancellation, speeded naming, etc.) makes sure that the phonological effects found in fact occur before lexical access (if lexical access occurs at all).

7. Note that in any case this represents some departure from model (1) since as stated in the diagram above, lexical access would be achieved by the phonologically recoded string only. Hence, one would expect many errors to occur with pseudohomophones. Though the number of errors is usually higher with pseudohomophones as compared with ordinary pseudowords, they are nevertheless rather rare (usually not more than 7-9%). However, more sophisticated versions of model (1) are conceivable where lexical access is not 'visually blind', that is, where visual information is not lost after phonological recoding. The trouble with such modelling, however, is the difficulty of distinguishing revised models from the family of dual access models (type 3).

8. This result is to my knowledge the first which can be interpreted as evidence that even in lexical decision, phonological effects do not prove that phonological recoding occurs prelexically. The point is stated explicitly in Gfroerer et al. 1984. In Günther et al. 1984, we argued that the lexical decision task is in fact rather unsuited for the investigation of "early" processes in word recognition.

9. Note that for pronouncing pseudowords, it is not appropriate to speak of a prelexical route since pseudowords have no lexical entries. Nevertheless, the usually used unqualified pseudoword argument holds that it is in fact the GPC road in models (1) and (3) which is used in order to achieve not only lexical access (and the pronunciation) for words, but also the pronunciation for nonwords.

10. Some evidence of this sort is provided in Günther & Greese (1985). Following an argument of Henderson (1982:159), the naming latency for lexical hermits was compared with non-hermits. Hermits are defined as words without lexical neighbours, i.e. words where the exchange of one single letter does not lead to another word. Also, pseudoword hermits can be defined as pronounceable letters strings which are not words and cannot be changed into a word by exchanging one letter. In the experiments, there was a significantly slower naming latency for

hermits in both words and pseudowords which cannot be explained by GPC theory since all stimuli were orthographically regular.

11. In the literature written in English, these patients are usually called dyslexics. German research follows the useful convention of using the prefix *a-* for acquired (language) pathologies and the prefix *dys-* for developmental ones. Hence, following this convention, surface dyslexics would be children who show a certain syndrome concerning disturbances in their learning to read, whereas surface alexics would be people with acquired reading pathologies of the same sort.

12. Besides these cases (deep alexia, phonological alexia and letter-by-letter reading), also surface dyslexia is sometimes cited as bearing evidence for the existence of a nonlexical GPC procedure in reading. This argument is rebutted quite clearly by Henderson 1982: 117ff.

13. For a directly opposite view, see Allport & Funnell 1981.

14. Notice that I do not deny the possibility that this conception, which is shared by most present theories of short term memory, might be true. Nevertheless, the empirical evidence for it is not at all compelling, even if theoretical arguments were.

BIBLIOGRAPHY

- Allport, A. & Funnell, E. 1981. "Components of the mental lexicon." In: H.C. Longuet-Higgins, J. Lyons & D.E. Broadbent (Eds.): *The psychological mechanisms of language*. London: The Royal Society, 397-410.
- Baron, J. & Thurston, I. 1973. "An analysis of the word superiority effect," *Cognitive psychology* 4: 207-228.
- Bierwisch, M. 1972. "Schriftsprache und Phonologie," *Probleme und Ergebnisse der Psychologie* 43: 21-44.
- Coltheart, M. 1978. "Lexical access in simple reading tasks." In: G. Underwood (Ed.): *Strategies of human information processing*. London: AP, 151-216.
- Coltheart, M., Davelaar, E., Jonasson, J.T. & Besner, D. 1977. "Access to the internal lexicon." In: S. Dornic (Ed.): *Attention and performance VI*. New York: AP, 535-563.
- Coltheart, M., Besner, D., Jonasson, J.T. & Davelaar, E. 1979. "Phonological encoding in the lexical decision task," *Quarterly journal of experimental psychology* 31: 489-507.
- Coltheart, M., Patterson, K. & Marshall, J. (Eds.) 1980. *Deep dyslexia*. London: Routledge & Kegan Paul.
- Coltheart, M., Patterson, K. & Marshall, J. (Eds.) 1985. *Surface dyslexia*. London: Lawrence Erlbaum.

- Conrad, R. 1964. "Acoustic confusion in immediate memory," *British journal of psychology* 55: 75-84.
- Coulmas, F. 1981. *Über Schrift*. Frankfurt: Suhrkamp.
- Davelaar, E., Coltheart, M., Besner, D. & Jonasson, J.T. 1978. "Phonological recoding and lexical access," *Memory & cognition* 6: 391-402.
- Gelb, I. 1963. *A study of writing*. Second edition (first 1958). Chicago:UP.
- Gfroerer, S., Günther, H. & Weiss, L. 1984. "Zur Frage des visuellen und phonologischen Zugangs zum mentalen Lexikon," *Phonetica* 41: 41-49.
- Gibson, E., Pick, A., Osser, H. & Hammond, M. 1962. "The role of grapheme-phoneme-correspondence in the perception of words," *American journal of psychology* 75: 554-570.
- Gibson, E., Shurcliff, A. & Yonas, A. 1970. "Utilization of spelling patterns by deaf and hearing subjects." In: H. Levin & J. Williams (Eds.): *Basic studies in reading*. New York: Basic Books, 57-73.
- Glushko, R.B. 1981. "Principles for pronouncing print: The psychology of phonography." In: C.A. Perfetti & A. Lesgold (Eds.): *Interactive processes in reading*. Hillsdale: Erlbaum, 81-98.
- Gough, P.B. 1972. "One second of reading." In: J. Kavanagh & I. Mattingly (eds.): *Language by ear and by eye*. Cambridge: MIT Press, 331-358.
- Günther, H. 1981a. "Das Prinzip der Alphabetschrift begreifen lernen: Einige Thesen zu einem fragwürdigen Konzept," *Forschungsberichte des Instituts für Phonetik und Sprachliche Kommunikation der Universität München (FIPKM)* 14: 53-68.
- Günther, H. 1981b. "Den Esel gemeint, aber den Sack geschlagen — Zu E. Scheerers Kritik an den Thesen zur Alphabetschrift," *FIPKM* 14: 103-120.
- Günther, H. 1982. Review of Coulmas 1981, *Zeitschrift für Sprachwissenschaft* 1, 301-304.
- Günther, H. 1983a. "Charakteristika von schriftlicher Sprache und Kommunikation." In: K.B. Günther & H. Günther (Eds.): *Schrift, Schreiben, Schriftlichkeit*. Tübingen: Niemeyer, 17-39.
- Günther, H. 1983b. "Studien zur visuellen Worterkennung," *FIPKM* 18: 1-190.
- Günther, H. 1984. "The priority of spoken language — On the invalid use of a valid concept." In: W.U. Dressler, O.E. Pfeiffer & J.R. Rennison (Eds.): *Discussion papers from the 5th international phonology meeting, Eisenstadt 1984*. Wien: WLG Beiheft 3: 103-107.

- Günther, H., Gfroerer, S. & Weiss, L. 1984. "Inflection, frequency, and the word superiority effect," *Psychological research*, 261-281.
- Günther, H. & Greese, B. 1985. "Lexical hermits and the pronunciation of visually presented words," *FIPKM* 21: 25-52.
- Günther, K.B. 1982. *Schriftsprache bei hör- und sprachgeschädigten Kindern*. Heidelberg: Groos.
- Haas, W. 1970. *Phono-graphic translation*. Manchester: UP.
- Hardyck, C.D. & Petrinovich, L.F. 1970. "Subvocal speech and comprehension level as a function of the difficulty of reading material," *Journal of verbal learning and verbal behavior* 9: 647-652.
- Henderson, L. 1982. *Orthography and word recognition in reading*. London: AP.
- Henderson, L. 1985. "Issues in the modelling of pronunciation assembly." In: M. Coltheart, K. Patterson & J. Marshall (Eds.), 459-508.
- Hörmann, H. 1983. "On the difficulties of using the concept of a dictionary — and the impossibility of not using it." In: G. Rickheit & M. Bock (Eds.): *Psycholinguistic studies in language processing*. Berlin: De Gruyter, 3-16.
- Huey, E.B. 1908. *The psychology and pedagogy of reading*. New York: Macmillan (Reprint 1968: MIT Press).
- Humphreys, G., Evett, L. & Taylor, D. 1982. "Automatic phonological priming in visual word recognition," *Memory & cognition* 10: 576-590.
- Liberman, A., Cooper, F.S., Shankweiler, D.P. & Studdert-Kennedy, M. 1967. "Perception of the speech code," *Psychological review* 74: 431-461.
- McCusker, L.X., Hillinger, M. L. & Bias, R.G. 1981. "Phonological recoding and reading," *Psychological bulletin* 89: 217-245.
- McQuade, D.V. 1981. "Variable reliance on phonological information in visual word recognition," *Language & speech* 24: 99-109.
- Messing, J., Günther, H. & Kegel, G. 1980. "Zur Fehleranalyse abweichenden Sprachverhaltens," *International review of applied linguistics in language teaching (IRAL)* 18: 326-334. Again in *FIPKM* 13: 7-18.
- Moraes, J., Cary, L., Alegria, J. & Bertelson, P. 1979. "Does awareness of speech as a sequence of phones arise spontaneously?" *Cognition* 7: 323-332.
- Nolan, S.D., Tanenhaus, M.K., & Seidenberg, M.S. 1981. "Multiple code activation in word recognition: Evidence from rhyme monitoring." Technical Report No. 204. Center for the Study of Reading, University of Illinois.

- Parkin, A. 1982. "Phonology recoding in lexical decision: Effects of spelling-to-sound regularity depend on how regularity is defined," *Memory & cognition* 10: 43-53.
- Patterson, K. 1981. "Neuropsychological approaches to the study of reading," *British journal of psychology* 72: 151-174.
- Rubenstein, H., Lewis, S. & Rubenstein, M.A, 1971. "Evidence for phonemic recoding in visual word recognition," *Journal of verbal learning and verbal behavior* 10: 645-657.
- Scheerer, E. 1978. "Probleme und Ergebnisse der experimentellen Leseorschung," *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie* 10: 347-364. (Again in Günther & Günther 1983, 88-103.)
- Scheerer, E. 1983. "Probleme und Ergebnisse der experimentellen Leseorschung - 5 Jahre später." In: Günther & Günther (Eds.), 105-130.
- Seidenberg, M.S. & Tanenhaus, M.K. 1979. "Orthographic effects on rhyming," *Journal of experimental psychology: Human learning and memory* 5: 546-554.
- Spoehr, K. 1978. "Phonological encoding in visual word recognition," *Journal of verbal learning and verbal behavior* 17: 127-141.
- Stanovich, K.E. & Bauer, D.W. 1978. "Experiments on the spelling-to-sound regularity effect in word recognition," *Memory & cognition* 6: 410-416.
- Steinberg, D.D. & Harper, H. 1983. "Teaching written language as a first language to a deaf boy." In: F. Coulmas & K. Ehlich (Eds.): *Writing in focus*. Berlin-New York: Mouton, 327-354.
- Tillmann, H.G. 1980. (with P. Mansell). *Phonetik*. Stuttgart: Klett-Cotta.
- Underwood, G. & Twaites, S. 1982. "Automatic phonological coding of unattended printed words," *Memory & cognition* 10: 434-442.
- Venezky, R. 1970. *The structure of English orthography*. Den Haag: Mouton.
- Weiss, L., Günther, H. & Gfroerer, S. 1984. "Orthographische Wohlgeformtheit und Lexikalität als unabhängige Faktoren des Wortüberlegenheitseffekts," *Sprache & Kognition* 3: 231-241.

ORTHOGRAPHY AS A VARIABLE IN PSYCHOLINGUISTIC EXPERIMENTS

Bruce L. Derwing and Maureen L. Dow

1. *Introduction*

By definition, every literate speaker of a language has acquired a working knowledge of at least one orthography, i.e., of some more-or-less standardized set of norms for representing a language in written form. Linguists have long recognized the “nuisance” character of this kind of knowledge, particularly in the sense that it tended to confuse or confound the linguistically untrained observer’s perceptions of the actual pronunciations of the words in his language (and was thus something which had to be systematically “overcome” as part and parcel of learning to view language in its “proper,” i.e., linguistic or speech-based, perspective). Though trained in this particular perspective ourselves, we have more recently been coming around to the view that much of what has often passed previously for investigation of the human *faculté de langage* is in fact rather heavily contaminated by many factors arising out of the process of getting a general education (including, in addition to knowledge of orthography, many other facets of prescriptive grammar, such as the ability to make judgments of syntactic well-formedness or “grammaticality”).¹ This paper reports on those aspects of our own psycholinguistic investigations which have shown orthography to be a significant experimental variable, including specific attempts to assess both the circumstances and extent of its influence. The importance of this line of research lies in the fact that, to the extent that knowledge of spelling (or any other aspect of world knowledge) influences a subject’s judgments or intuitions in an experimental task, the results of such an experiment cannot be validly interpreted without taking this factor into account.

2. Some examples of orthographic effects

Our first major encounter with the orthographic variable occurred in conjunction with a series of experiments intended originally to assess the role of phonetic and semantic similarity in the task of morpheme recognition.² On the basis of a pseudo-etymological “comes-from” task to assess morpheme identity, this work indicated that word pairs were readily found to contain a common morpheme only when a semantic similarity rating of at least 50% was (independently) assigned to the words and a phonetic similarity rating of about 30% or more. Some notable exceptions, however, were words like *breakfast*, *handkerchief* and *cupboard*, whose constituent root morphemes, *break*, *hand* and *cup* were identified far more readily than would have been predicted on the basis of sound and meaning similarities alone. This result was confirmed in later work involving different experimental techniques, such as a more straightforward morpheme identification task, in which, after a brief training session with a few simple examples (e.g., *cat/s*, *play/er*, *foot/ball*, *smok/y*, *sad/ly*, etc.), subjects were asked to divide (orthographically presented) English words into their constituent “meaning units.” Not surprisingly, this last test seemed particularly prone to orthographic effects, bringing out as it did not only the same old anomalies, but also a substantial number of new subjects who found the (historical) *lace* in *necklace* (80%), a *slip* in *slipper* (53%) and even, for some, a *bash* in *bashful* (27%) and a *hide* in *hideous* (25%)!

By pure coincidence, attention was diverted from the study of morphology in this work by the appearance of an article by Vitz & Winkler (1973), which indirectly suggested some interesting applications for phonology. The paper itself was a follow-up of an earlier study by Nelson & Nelson (1970), who found that both phonetic and orthographic similarity between visually presented pairs of words could influence judgments of “similarity of sound.” V&W attempted to quantify a set of predictions of such judgments in terms of a (non-orthographic) phoneme model, with generally rather encouraging results (though with somewhat less success on N&N’s particular word set). Derwing (1976) confirmed V&W’s main results with the word set of the morphological study outlined above, but also noted that an “orthographic index,” calculated according to the same logic as V&W’s “phonemic index,” yielded a set of predictions that were, overall, almost as good (both correlations over .80). In an attempt to assess the relative contributions of orthographic (graphemic) and phonemic

(sound) similarity in this type of task, Derwing used partial correlations to factor out the independent effects of these two variables and found that the latter was clearly more important than the former, though both were significant.

The generally high predictability of "sound similarity" judgments in terms of a simple phoneme-count model, therefore, together with a relatively low estimate of extraneous orthographic effects, next gave rise to a series of experiments designed to clarify some rather long-standing controversies of phonological/phonemic theory, including, in particular, the following three: (1) the segmental status of English vowel nuclei (notably such diphthongs as appear in the words *pain/pane* or *soul/sole*): should they be represented as single vowel phonemes (as implied by the transcriptions /pen/ and /sol/ cf. Pike 1947 and Fries 1948) or as double segments (i.e., vowel plus glide, as in /peyn/ and /sowl/, following the influential recommendations of Trager & Smith 1957); (2) the phonemic quality of English stops after syllable-initial /s/, as in *spill, still* and *skill*: in the absence of any voiceless/voiced contrasts in this environment, should these voiceless, unaspirated stops be analyzed as /p,t,k/ (in accord with the usual linguistic treatment and, normally, with the standard orthography, as well; cf. Hockett 1958, Gleason 1961, etc.), as /b,d,g/ (in accord with the results of phonetic "gating" experiments; cf. Lotz, *et al.* 1960, Welz 1982), or as the (archi)phonemes /P,T,K/, distinct from both sets (cf. Trubetzkoy 1939); and (3) the (psychologically) best analysis of the English phone [ŋ]: do speakers treat it as the single phoneme /ŋ/ (as in most classical, pre-generative analyses), or rather as the (abstract) sequence /ng/ (in accord with the orthography and as advocated by Chomsky & Halle 1968, and, in a particularly famous case, by Sapir 1925)? The logic behind all three categories of "string similarity" experiments was essentially the same. In the first case, the phonemicizations /peyn/ 'pain' vs. /pen/ 'pen' predict a higher degree of judged similarity (three shared phonemes out of four, or .75) than do the phonemicizations /pen/ 'pain' vs. /pen/ 'pen' (two out of three, or .67). In the second case, the phonemicizations /spil/, /sbil/ and /sPil/ for 'spill' give differential similarity predictions with such words as /pil/ 'pill' and /bil/ 'bill' (viz., .75 and .50, respectively, for /spil/, .50 and .75 for /sbil/, .50 and .50 for /sPil/, etc.). And in the third case, of course, the representations /sŋ/ vs. /sing/ for 'sing' make differential predictions about the distance of this word from another such as /sin/ 'sin' (viz., .67 vs. .75, respectively).

Unfortunately, only the first problem has yielded anything like a con-

sistent set of results in this type of task: all English vowel nuclei, it seems, are best represented as single phonemic segments.³ In the case of the stops, while our data generally support the voiceless analysis, this is only consistently true when the analysis is directly supported by the spelling (as in *pill-spill* vs. *bill-spill* and *kill-skill* vs. *gill-skill*). The results are less clear if the spelling relationships in the word-pairs used are even marginally less straightforward (as in the comparisons *cull-skull* vs. *gull-skull*) and the effect disappears entirely in the case of *cot-squat* vs. *got-squat*.⁴ In our attempt to control for orthographic influence through the use of nonsense words (e.g., *pif-spif* vs. *bif-spif*), the results still support the voiceless analysis, but only marginally so.⁵ (See Derwing & Nearey (1986) for a summary of the data utilized here.) Thus, even though our data never argue in favor of either of the competing analyses, we can not yet exclude the possibility that our subjects' knowledge of basic English spelling conventions (e.g., that *sb*, *sd* and *sg* are quite impossible spellings in syllable-initial environments) might well be the main determining factor here. As for the English phone [ŋ], the "string similarity" technique has so far failed completely to yield any clear preference for either /ŋ/ or /ng/ (and this despite the quite obvious orthographic bias in favor of the latter), whether due to inherent uncertainty or variability on the part of our subjects themselves or to inadequate sensitivity in our experimental technique. Discouraging as this may at first seem, it is nonetheless interesting that, given now the perspective of Smith 1982, even the available speech error data (once widely believed to be rather conclusive in favor of the /ng/ analysis) must now also be regarded as quite indeterminate on this point, as well.

Without a doubt, there are other long-standing controversies in English phonology that are likewise begging for resolution but are evidently not likely to find it in the near future, either because of indeterminacy or variability in subjects' own intuitions or because subjects use the traditional orthography as a guide in making the relevant judgments (cf. O'Neil 1972, Derwing 1973 and Skousen 1982, for example, on the treatment of the flaps in the classic *writer/rider* example). Read (1973) has brought this out in a particularly telling illustration of the English word-initial *tr* cluster, in which the "t"-sound is characteristically retroflexed and/or affricated. Significantly, whereas literate adults invariably interpret this sequence as a /tr/, in accord with the orthography (cf. the words *tree* and *truck*), Read's preliterate child subjects commonly preferred the spelling *chr*, implying the analysis /čr/. So just what is going on here? Does the language learner's

interpretation of this sequence undergo a change as he matures, or is the adult's judgment overwhelmingly influenced by the spelling — or is this merely two different ways of saying essentially the same thing (cf. Ehri & Wilce, 1980)?

3. *The orthographic variable as a specific object of study*

The considerations raised above highlight the necessity to achieve a better understanding of the orthographic variable itself, both with respect to the nature and extent of its influence in psycholinguistic experiments and to the variability of this influence as a function of particular experimental task demands. The best experimental work we know of dealing with the first of these questions is that of Linnea Ehri of the University of California at Davis. She has shown, for example, not only that there is a correlation between the number of letters used to spell certain words and the number of speech sounds that speakers segment them into (cf. *pitch* vs. *rich* and *badge* vs. *page*), but also that, in controlled situations using nonsense words, these "extra segments" are perceived almost exclusively by those subjects who learn the longer spellings (cf. *zitch* vs. *zich* and *tadge* vs. *taj*). She has also discovered that this kind of orthographic influence on phonological judgments is not completely "blind": letter combinations such as *sh* and *ch* are generally understood to represent single, not double, phonological units; "silent" letters are not tabulated at all (e.g., the final letters of the words *home* and *comb*); and (consistent with our own findings as noted in Section 2 above) vowels tend to be interpreted as single segments regardless of their spelling (cf. *own* vs. *old* and their nonsense counterparts *drowl* vs. *drol*).⁶ (See Ehri, in press, for a good summary of her research in this area.)

Finally, two related questions concern the possible interaction of the orthographic variable with the experimental task. Are some tasks inherently more heavily influenced by knowledge of orthography than others, and, given a particular experimental task, is it possible to weaken (perhaps even effectively eliminate) the effects of spelling, say, by so simple a device as changing from a written (or visual) to a spoken (or oral) mode of presentation? As already noted in Section 2 above, the results reported in Derwing & Nearey (1986) suggest that the answer to the first of these questions must be in the affirmative; other work, however, has suggested that the answer to the second question may be in the negative (e.g., the early

“string similarity” study reported in Derwing (1976), which yielded virtually identical results for both visual and oral groups). Dow (1981) was designed to focus on these two specific issues and to clarify the answers a good deal further.

The two tasks selected for this study were a rhyme task (Experiment 1) and a “segment count” task, i.e., a task that involved a straightforward count of the number of “speech sounds” in a particular set of words (Experiment 2), with presentation mode included as a variable in both experiments. The stimulus words were presented orally to all subjects, in order to insure that there was no ambiguity as to intended pronunciation, while for half of the subjects in each group the words were also presented visually (written on the answer sheets). Thus the effect of both type of task and presentation mode was tested.

Experiment 1: Rhyme task

A rhyme task seems well-motivated when one reviews phonological studies requiring similarity of sound judgments, where rhyme has consistently been found to be an important parameter. Nelson and Nelson (1970) found that words with the last two phonemes in common (out of three) were rated as more similar than those with the first two, the first, or the last phoneme in common. They termed this the “rhyme effect.” In the Vitz and Winkler (1973) study, post-experimental interviews were held concerning what subjects felt was influencing their similarity of sound judgments, and twenty-five percent of those reports related to rhyme. Derwing and Nearey (1986) also found a rhyme-related effect, namely, that differences between consonants in final position had a greater effect on sound similarity judgments than differences between initial consonants.

Rhyme in the English language is ideal for the investigation of orthography in phonological judgments for a very practical reason, as well. The English orthography has been much maligned because of its many ways of spelling the same sound, but in this study that aspect proves very useful. One example of this is the alternate spellings of the phonemic sequence /a^yt/ (e.g., *light* vs. *kite*). If the spelling of the test word influences the choice of rhyme, then in all likelihood the rhyme provided for *tight* would be a word like *light* rather than a word like *kite*.

In this experiment, 40 high school students were presented with words illustrating spelling alternations for rhymes and were asked to provide a rhyming word for each word given. Responses were written, to avoid any

possible ambiguity in scoring due to homophony. In one condition, the subjects only heard the words and did not see them; for the other group of subjects, the words were presented both orally and visually. There were 120 words made up of six words from each of 20 English rhymes with alternate spellings (chosen from a larger list). All were common words with a relatively large number of possible rhymes. Scoring was a simple same/different judgment, reflecting whether the rhyme-element of the word provided by the student was spelled the same as or differently from the test word.

If knowledge of orthography is an important factor in such an elicitation task, one would expect responses to be non-random, with a tendency for the rhyme-element of the word provided by the students to be spelled the same as that in the test word. If, alternatively, orthographic effects vary as a function of presentation mode, one would expect a difference to emerge between the oral and written groups. The third possibility would be that orthography does not play any role at all in such a task, and responses of both groups would be random with respect to the spelling of the rhyme-element of the words presented. In such a case, any small trends found might be due only to such extraneous factors as word frequency, idiosyncratic spelling, individual preferences, etc.

There was a clear effect found for visual presentation. Not only was the mean score of same spellings with written presentation significantly higher than with oral presentation (74.9 vs 64.9), the latter was not significantly different from chance. This would suggest that the basis for choice of rhyme with the oral presentation was solely the sound of the test word. The students in the written presentation group, on the other hand, must have been somewhat influenced by the spelling of the test word in their choice of rhyme. Post-experimental interviews supported this conclusion: a number of students in the written group claimed they had noticed that many of their rhymes matched the spellings, whereas none of the oral group seemed to have made any connection between the task and the spelling of the words.

Analyses were also carried out to determine whether there were preferred spellings for each of the individual rhyme groups. Although this turned out to be so in 11 of the 20 cases, the preferred spellings did not seem to be related to any simple frequency measures at a level greater than chance. We also looked at the possibility that spellings might be preferred which conformed to some general spelling rule, such as the rule that "a VCe sequence triggers the 'long' quality of the first vowel." However, our results were not suggestive of influence from this rule, since the six rhyme groups

for which the rule is relevant were clearly split in preference: in three of the cases, the “silent e” spellings were preferred, but in the other three cases the alternate, *ad hoc* spellings were dominant.

Another possibility might be that there are subgroups of preferred spellings, with separate explanations for each preference. For example, educated or prescriptive spellings might have been preferred for some of the rhyme groups. Thus, given the prescriptive bias against such spellings as *nite* and *lite*, high school students asked to participate in a linguistic experiment might be inclined to avoid the *-ite* spelling in general in favor of words spelled with *-ight*. Or perhaps some forms might, for one reason or another, be simply “harder to spell.” One example of this seems to be the *-our* form for the /a^wr/ group. The more predictable and “easier” spelling for that rhyme is apparently the *-ower* one, perhaps because syllabic /r/’s are so commonly represented orthographically in English as *-er*. Some support for this explanation might be found in English spelling errors and in experiments that test for the spelling of nonsense words.

To test the suggestion that *-ower* is the productive form, a small supplementary study involving nonsense words was run. Eight adult native English speakers were asked to spell eleven words of the /a^wr/ rhyme group. Five of these were real words (*scour*, *glower*, *bower*, *sour*, *flour/flower*) and six were nonsense words (/sna^wr/, /sta^wr/, /kla^wr/, /ma^wr/, /fa^wr/, /ja^wr/). The real words were spelled correctly by all subjects, with *flour* the preferred form for the homophonic pair. For the nonsense words, the majority of the spellings were *-ower* (26 as opposed to 14 of the *-our* form). Note, however, that two of the nonsense words (/sna^wr/ and /ma^wr/) would have had English homographs with a different pronunciation if spelled *-ower* (i.e., *snower*, *mower*) and this may explain why that form was rejected in these cases. Interestingly, the spellings for these two nonsense words were split between *-our* and *-auer* (on analogy, perhaps, to the trade name *Bauer*). Subjects tended to try the *-ower* spelling first in these cases and then changed it, apparently to avoid the ambiguity. These results thus provide some support for the claim that *-ower* is the productive form for this rhyme group.

The varied suggestions to explain the preferred spellings point out that there are probably influences on phonological judgments quite separate from sound perceptions *per se*, and orthography may be only one of these. In order to understand the language knowledge of native speakers and to make accurate predictions about performance in experimental situations, a full specification of these influences is necessary. A careful re-evaluation of

traditional phonological analyses is also necessary to insure that native speaker competence has been accurately represented. This, and the question of whether the effect of presentation mode is generalizable to all tasks tapping phonological knowledge, was the motivation for the second experiment.

Experiment 2: "Segment count" task

The emphasis in phonological and phonetic experimentation has generally been on tasks involving comparison judgments of similarity of sound, discriminations between tokens, and classifications of tokens into phonemic categories. The latter two tasks restrict the scope to the level of single phonemes. The comparison judgments allow the use of whole words and syllables, but only indirectly tap phonological knowledge, and actual phonemicizations can only be guessed at. Linguistically naive subjects have been preferred for such experiments in order to avoid undue influence from knowledge of a particular theory or the expected results. It has thus been virtually impossible to elicit phonemicizations directly, since some familiarity with phonetic transcription and phonemic theory would be required for such a task. Some investigators have resorted to training their subjects, but in such cases the question again arises as to whether the subjects can any longer be considered naive, since the experimenter has imposed both his phonetic transcriptions and his segmentation biases on them.

The problem, therefore, becomes one of devising a task that can access the phonological knowledge of naive subjects in a more direct way than through comparisons or classifications. One straightforward suggestion that seemed worth trying was to ask subjects simply to write down the number of "speech sounds" in token words. No definition of the phrase "speech sounds" was provided in the initial phase of the experiment, leaving the interpretation of the term up to the subjects' own devices. The motivation for this was to prevent biasing of responses in the direction of any *a priori* notion of phonetic segmentation. This resulted in the use of a variety of different strategies by different sub-groups of subjects on this task. Of the 40 high school students tested, therefore, less than a third ($n = 13$) appeared to do their counting on a strictly segmental (i.e., phonemic) basis; these subjects performed comparably to a separate group of 20 students in the second phase of the experiment, who were provided with a few sample segmentations as a guide (viz., *cat* = 3 sounds, *debt* = 3, *creep* = 4, *laugh* = 3, *ghost* = 4).

One key question of interest was whether presentation mode would be an important effect in this task, as it was in Experiment 1, so both visual and oral presentation modes were again used. If the segmentation criterion was phonemic, or something approximating the phonemic notion of the segment, then the responses for this word should have been in the mid-range (2-5). If the criterion was letter counting, or if the number of letters heavily influenced the judgments, then the numbers given should have been slightly higher than for a phonemic count (2-7) and responses would be correlated to a grapheme, rather than a phoneme, index. In addition, on analogy to Experiment 1, letter-counting was expected to be more prevalent in the written presentation group than the oral group. In contrast to the rhyme experiment, there was no significant difference between performance of the oral and written presentation groups in either phase of the speech sounds task. It was therefore concluded that the effect of presentation mode was not significant in this experiment.

As expected, there was a strong correlation of the number of speech sounds given to the phoneme index ($r = .92$). The correlation of grapheme index with speech sounds was lower ($r = .79$), but still significant at the .01 level. To determine the relative importance of phonology and orthography in these correlations, partial correlations were calculated in the manner of the Derwing (1976) study already noted in Section 2 above. The correlation of phoneme count with speech sounds controlling for orthography was $r = .86$, while the correlation of letter or grapheme count with speech sound controlling for phonology was $r = .57$. All of these correlations were significant at the .01 level. (The 13 segment-oriented students from the first phase of the speech sounds task showed the same pattern, with the same level of significance.) Overall, therefore, phonology once again appeared to be more important than orthography in this task (but see n. 6), and this lack of influence of orthography extended to the nonsignificance of presentation mode, as well (i.e., seeing the spelling of the words did not seem to influence the answers given for number of speech sounds).

There were, however, a small set of words that had means for the number of speech sounds that were quite different from the theoretical phoneme count. The words *cue*, *pew*, and *view* were problems, as were *cower*, *bower*, and *glower*, as well as *quote* and *mute*. The phonemicizations that were the basis of the phoneme index used in this task counted all vowel nuclei as single segments, thereby including the post-vocalic glide elements of diphthongs with the vowel, but pre-vocalic glides were rep-

resented separately (as in the words /yul/ 'Yule' and /wet/ 'wait'). The reason for this method of segmentation was that it yielded predictions as to the number of speech sounds in a word that best reflected native speaker responses in previous experiments. The exceptions listed above obviously had been interpreted in ways that were not in accord with this analysis.

One group of exceptions included words that all contained pre-vocalic glides in post-consonantal environments (*pew, cue, view*). There was a strong tendency for subjects to claim that there were two, rather than three, speech sounds in these words, so that the glide element was evidently perceived either as part of the initial consonant, or of the vowel. (The two other exceptions that are relevant there are the words *mute* and *quote*, which subjects tended to say consisted of three, rather than four, speech sounds.)

The only other exceptions also formed a homogeneous class in that they all belonged to the same rhyme group, viz., the words *glower, cower* and *bower*. Quite unexpectedly, all three of these words were consistently rated differently from their presumed rhymes, *sour, scour* and *dour*. Evidently, rather than counting the /a^wr/ part of the -ower words as two segments, our subjects rated it as three, unlike the words with the -our spellings. At least two possible explanations are suggested for this phenomenon. First, perhaps the presence or absence of the orthographic *w* can, under some circumstances, affect the perception of the glide element of a diphthong. Thus, in these examples, when a word was spelled without the *w*, the vowel and the glide were perceived as a single segment, but if a word was spelled with a *w*, the glide might have been perceived as a distinct segment. An alternative explanation for the apparent differences in perception between the two spellings involves the perception of syllabic /r/. Perhaps in words where the /r/ is represented orthographically as -er, it may be perceived as /ər/ (i.e., as two segments), and elsewhere as a single segment. This, too, would explain why the -ower words were all rated as having one more speech sound than the rhyming -our words.

The examples that have been discussed above concern the possible misrepresentation of native speaker perceptions by traditional phonemic analyses. Although not the original motivation for this study, the delineation of such discrepancies between phonological analyses and speaker perceptions should help explain some results that until now have been rather enigmatic. Instead of casually blaming the "possible confounding influence of orthography" for results that do not fit with established theories and pre-

dictions, phonologists need to examine their assumptions about speaker knowledge more closely and to conduct whatever experiments are required to evaluate all such confounding variables.

However, it seems clear that orthography can be a very important influence on phonological judgments, depending on the type of task. In the rhyme task, presentation mode was a significant effect, but in the speech-sound task it apparently was not. This suggests that knowledge of orthography might be differentially utilized, depending on the type of task, and with variable influence on phonological judgments depending on the degree or kind of phonographemic divergence involved. In ambiguous cases, where there are no morphological cues, orthography in the form of "careful pronunciations" might well be the deciding factor in making some phonological judgments. A possible example of this kind of ambiguity can be illustrated in the following hypothetical dialogue:

Q: Which do you want, the ladder or the pail?

A: The ladder/latter.⁷

Of particular interest to phonology are those words that have had underlying representations postulated largely or exclusively on the basis of presumed morphological relationships to other words (e.g., *sign/signal*). The question of whether or not there is a /g/ in a speaker's internalized lexical entry for the morpheme *sign* is a central one for phonological theory and bears on the general issue of abstractness in phonology (since no overt phonetic [g] ever appears in this word). But even if we were able to establish clearly that the answer to this particular question was in the affirmative (at least for most speakers), we would still be faced with the issue of whether it was indeed morphological relationships or simply knowledge of the orthography that was ultimately responsible for this.⁸ There is no reason to believe that questions of phonological representation or segmentation might not be resolved by speakers on the basis of whatever information, "linguistic" or otherwise, that is available to them, and issues of language learning, knowledge and use can not so easily be distinguished from more general issues of learning, memory and cognition. Language touches on almost all of a speaker's life experiences and there are probably no "pure" linguistic judgments that can be accessed in isolation from the effects of these far-ranging influences.

NOTES

1. See Derwing & Nearey (in preparation) for an elaboration of this point.
2. See Derwing (1973:122-126) for a discussion of the theoretical background for these studies and Derwing (1976) for a description of the first round of experiments. The "meaning unit" study noted below is described in Derwing & Nearey (1986).
3. After consonants, this may even include vowels with *pre-vocalic* glides, as in *pew*, *mute* and *quote*, as noted in Section 3 below.
4. But see the discussion of the pre-vocalic glide phenomenon in Section 3 for evidence of another possible contaminating factor in this last comparison.
5. Moreover, as Frith (1980) has noted, literate speakers may well be able to supply implicit spellings even for novel nonsense words. There is also some evidence for this in Derwing & Nearey 1981.
6. It should be noted here that Derwing's "grapheme index," as discussed in Section 2 above, was *not* sensitized in this way, with the result that any conclusions based upon it must be subject to reconsideration.
7. Thanks to Ron Smyth for this example.
8. Sapir (1925) dismisses the orthography issue out of hand in favor of more "linguistic" considerations, but gives no compelling reason for doing so.

BIBLIOGRAPHY

- Chomsky, N. & Halle, M. 1968. *The sound pattern of English*. New York: Harper & Row.
- Derwing, B.L. 1973. *Transformational grammar as a theory of language acquisition*. Cambridge: Cambridge University Press.
- Derwing, B.L. 1976. "Morpheme recognition and the learning of rules for derivational morphology," *Canadian journal of linguistics* 21: 38-66.
- Derwing, B.L. & Baker, W.J. 1979. "Recent research on the acquisition of English morphology." In: P. Fletcher & M. Garman (Eds.), *Language acquisition: Studies in first language development*. Cambridge: Cambridge University Press.
- Derwing, B.L. & Nearey, T.M. 1981. "On the perceptibility of subphonemic differences: The *tough-duck* experiment," *Journal of the Atlantic Provinces Linguistic Association* 3: 29-40.
- Derwing, B.L. & Nearey, T.M. (1986). "Experimental phonology at the University of Alberta." In: J.J. Ohala & J.J. Jaeger (Eds.), *Experimental phonology*. New York: Academic Press.
- Derwing, B.L. & Nearey, T.M. (In preparation). "On the empirical con-

- tent of grammatical theory."
- Dow, M.L. 1981. *On the role of orthography in experimental phonology*. Unpublished M.Sc. thesis, University of Alberta.
- Ehri, L.C. (In press). "Effects of printed language acquisition on speech." In: D. Olson, N.G. Torrance & A. Hildyard (Eds.), *Literacy, language and learning: The nature and consequences of reading and writing*. Cambridge: Cambridge University Press.
- Ehri, L.C. & Wilce, L.S. 1980. "The influence of orthography on readers' conceptualization of the phonemic structure of words," *Applied psycholinguistics* 1: 371-385.
- Fries, C.C. 1948. *Teaching and learning English as a foreign language*. Ann Arbor: University of Michigan Press.
- Frith, U. 1980. "Unexpected spelling problems." In: U. Frith (Ed.), *Cognitive processes in spelling*. London & New York: Academic Press.
- Gleason, H.A. 1961. *An introduction to descriptive linguistics*. New York: Holt, Rinehart & Winston.
- Hockett, C.F. 1958. *A course in modern linguistics*. New York: Macmillan.
- Lotz, J., Abramson, A.S., Gerstman, L.J., Ingemann, F., & Nemser, W.J. 1960. "The perception of English stops by speakers of English, Spanish, Hungarian, and Thai: A tape cutting experiment," *Language and speech* 3: 71-77.
- Nelson, D.L. & Nelson, L.D. 1970. "Rated acoustic (articulatory) similarity for word pairs varying in number and ordinal position of common letters," *Psychonomic Science* 19: 81-82.
- O'Neil, W. 1972. "Our collective phonological illusions: Young and old." In: J.F. Kavanagh & I.G. Mattingly (Eds.), *Language by ear and by eye*. Cambridge, Mass.: MIT Press.
- Pike, K.L. 1947. *Phonemics: A technique for reducing languages to writing*. Ann Arbor: University of Michigan Press.
- Read, C. 1973. "Children's judgments of phonetic similarities in relation to English spelling," *Language learning* 23: 17-38.
- Sapir, E. 1925. "Sound patterns in language," *Language* 1: 37-51.
- Skousen, R. 1982. "English spelling and phonemic representation," *Visible language* 16: 28-38.
- Trager, G.L. & Smith, H.L., Jr. 1957. *An outline of English structure*. Washington, D.C.: American Council of Learned Societies.
- Trubetzkoy, N. 1939/1969. *Principles of phonology*. Originally published in German (*Grundzüge der Phonologie*) as *Travaux du cercle linguistique*

- de Prague 7.* Translated by C.A.M. Baltaxe. Berkeley & Los Angeles: University of California Press.
- Vitz, P.C. & Winkler, B.S. 1973. "Predicting the judged 'similarity of sound' of English words," *Journal of verbal learning and verbal behavior* 12: 373-388.
- Welz, T.T. 1982. *Categorization and discrimination of word juncture in English*. Unpublished M.Sc. thesis, University of Alberta.

SPELLING AND DIALECT

Gunnel Melchers

In writing up this study I first of all wish to express my gratitude to Mr. Alec Stout of Barkland, Fair Isle, whose questions on dialect writing first set me thinking seriously about problems connected with the orthographic representation of non-standard dialect.

Introduction

Discussing some basic concepts in dialectology, a recent, widely used textbook on ‘varieties of English’ (O’Donnell & Todd) states that dialects — as opposed to languages — are “characteristically *spoken*”. This is indeed a truth with qualifications — there are countless examples of non-standard writing (private as well as literary) in most literate societies from various periods in history. The situation today is that we are experiencing changing attitudes to dialects and accents in that native speakers of non-standard dialects openly wish and are encouraged to retain their local speech or, at least, to acquire bidialectal competence. Social changes, a growing concern for environment and local identity, the ‘roots syndrome’, and a general boosting as a result of sociolinguistic research are no doubt some of the factors behind these new attitudes, this consciousness-raising. Consider, for example, Labov’s classical Martha’s Vineyard study, Trudgill’s “Accent, dialect, and the school”, and the investigation of language attitudes on the island of Gotland, Sweden, with the telling title of “Vi har inte lust att prata nån jäkla rikssvenska!” (“We don’t feel like speaking some bloody Standard Swedish!”). One result of this increasingly positive attitude to local, non-standard varieties is, in fact, a growing interest in dialect *writing*, found particularly in areas where bidialectism can be defined clearly, i.e. where there is a discrete choice between the standard and ‘traditional dialect’ (cf. Wells 1980:2-8).

A linguistic area of such a type is definitely found in the Shetland Isles,

the northernmost part of Britain. Most Shetlanders are competent in two varieties of English, generally referred to as 'Shetland' and 'English'. 'Shetland', the traditional dialect, can be described as a variety of Lowland Scottish with a substantial component of Scandinavian, particularly noticeable in the lexicon but also 'stamping' other levels of language such as phonetics/phonology and syntax. It should be pointed out that this component differs from the Scandinavian linguistic heritage in other parts of Britain (with the exception of Orkney and, possibly, Caithness), not only in size but also in structure and history. In Shetland a form of Scandinavian, known as 'Norn' (ON norroenn), was the first Germanic language to be spoken there, and it was the dominant language for about 500 years. If the islands had remained under Scandinavian rule, Shetland language today would probably be a variety rather like Faroese. 'English', the variety used in speaking to non-Shetlanders and in certain formal situations, including most classroom situations, can be characterised as a fairly general British standard as to grammar and lexicon but with a distinct Shetland accent. It differs, however, from 'Shetland' to some extent as to the phonological specifications of lexical items, i.e. of the lexical incidence of particular phonemes in particular words (cf. Wells 1980:5).

Bidialectal competence is now part of the policy of Shetland schools.

"Dialect should not be condemned but regarded as a valid form of speech appropriate to certain situations. If a clearly defined policy of bilingualism is cultivated by the school, pupils should automatically know when to use dialect and when Standard English."

(From an instruction given to teachers at The Anderson High School, Lerwick, by John J. Graham, former headmaster and also author of *The Shetland dictionary*. The instruction is quoted in full in Melchers 1983:6.)

Many teachers also encourage bidialectal competence in *writing*. Attending English lessons in various schools all over Shetland, I have witnessed how children have been told to write essays or dialogues, sometimes in 'English', sometimes in 'Shetland'. Although there are no fixed spelling conventions for the traditional-dialect (cf. further below), the children seemed to be able to produce their writing quite effortlessly.

In the spring of 1983 I carried out an investigation of language attitudes among pupils of The Anderson High School. One of the questions was: "Do you sometimes use dialect words and phrases in composition and written tests?" with the alternatives 'often', 'sometimes', 'rarely', and 'never'. Quite a few children did not like any of the alternatives but produced reac-

tions such as: "I use dialect if it is a Shetland-based piece of writing", "I never use dialect, but if it was a story about Shetland and the teacher told us to, I would", "Depends if you are meant to write in Shetland. Stupid question".

Shetland cannot pride itself on having a major writer of fiction like Orkney's George Mackay Brown. An amazing number of books, magazines, and journals are, however, produced on the islands, mostly dealing with aspects of Shetland life, and often containing a great deal of dialect writing in the form of poems, short stories, and diaries. One of the magazines runs a dialect crossword, and there is an annual competition for the best efforts in writing 'Shetland'.

There are even local publications in various areas on Shetland Mainland and on outlying islands. Fair Isle, 24 miles SW of Mainland, with a total area of 2 sq. miles and a population of 80, has its own "Fair Isle Times". Most issues contain a passage in the local dialect, written by Alec Stout, who has lived on Fair Isle all his life and is very much concerned about the preservation of his native speech.

This is a sample of his writing:

"Onywyne, Up-Helly-Aa duy kem roond, in we riggat wiz athin wir sunday suits in polished wir schön, in stöd raedy for Wullie Aervin's bus ti pick wis up. Dat mornan, da weddar wizna göd, in de snaw wiz tehwan, aa'thing wiz athin a slushy maess in hit wiz braw cauld. Da bus duly kem assythte, in we klem abörd, duy wir a braw lokk a fauk in um. We roykatt da toon at de back o' eelivan in de driver perkit da bus anunder da 'faurt' at da nort aend o' eelivan in de driver perkit da bus anunder da 'faurt' at da nort aend o' da street. Weel, da street wiz athin some aelt, aboot twartree inch o' weet slush, but hit niver dampened wir ego, hit jüst steepit wir bonnie sheeny schön till duy wir a dirty ego, hit jüst steepit wir bonnie sheeny schön till duy wir a dirty gruy colour, in wir feet kinda weet. Weel, we a' six plüttird trehw da toon aw duy, spikkan ta fauk dat we kent in leukan in da shop winda's. I mind wis sittin ida 'Auld Rock Cafe' aettin wir denner, watchin da fauk gain by, in huyan a göd lauch. Da corn a haet maet did wiz a lokk a göd, warmed up wir boedies in keepit da fraust bite fuy settin in. We spent da eenin wanderin aboot, in sittin ida bus. I tink we güd fur a cup a tea in suntin ta aet afore lyhtin up time, hit bein aboot sivin a'klokk."

(Translated into Standard English:

"Anyway, Up-Helly-Aa day [a Shetland festival in January] arrived, and we put on our Sunday suits and polished our shoes, and stood ready for Willie Irvin's bus to pick us up. That morning the weather wasn't good, and the snow was thawing, everything was in a slushy mess and it was quite cold. The bus duly came in sight, and we climbed aboard, there were quite a lot of people in it. We reached the town just after eleven, and the driver

parked the bus below the ‘fort’ at the north end of the street. Well, the street was in some kind of dirty mess, some inches of wet slush, but it never dampened our ‘ego’, it just soaked our beautiful shining shoes, until they were a dirty grey colour, and our feet sort of wet. Well, all six of us ‘squelched’ about in town all day, speaking to people we knew and looking in the shop windows. I remember us sitting in the ‘Old Rock Café’, eating our dinner, watching the people going by, and having a good laugh. The bit of hot food did us a lot of good, warmed up our bodies, and kept the frost-bite from setting in. We spent the evening wandering about, and sitting in the bus. I think we went for a cup of tea and something to eat before lighting-up time, it being about seven o’clock.”)

As Mr. Stout is one of my chief informants for a project I have been involved in for a number of years (“The Scandinavian Element in Shetland Dialect”, cf. Melchers 1983), I have often had occasion to discuss the problems connected with giving a true representation of a particular dialect. Fair Isle speech has a rather special and complicated phonology, not only differing from RP or a Scottish standard, but also from other Shetland accents, and the limitations of the ordinary English alphabet are only too obvious. Notice, for example, the use of the ö and ü symbols in the above passage.

Mr. Stout has, in fact, turned to me and my colleagues in the project for practical advice as to the spelling problem. This seems to me to be something of a linguist’s dream. Challenged by this immediate task, I have therefore chosen to present a case study of ‘writing in Fair Isle dialect’, which, on one level, is the aim of this study. In a wider perspective, the discussion should be relevant for non-standard writing in general, at least of varieties of English, from which language most of my examples will be taken. Various examples of ‘regional spellings’ and suggestions for spelling reforms will be discussed, as well as possible teaching implications.

The study proper is restricted to phonetic/phonological representation, but an overview of dialect writing in general with some discussion of relevant terminology will first be given.

‘Writing in dialect’ as a field of study

A distinction is sometimes drawn between ‘literary dialect’, taken to mean the use of non-standard features in the language of more or less well-established writers of fiction, usually restricted to the dialogue,¹ and ‘dialect literature’, defined as attempts by native speakers of a dialect to give an orthographic representation of their speech in poetry or fictional prose (cf., e.g. Wakelin, *English dialects* and *Discovering country dialects*). It seems

to me, however, that such a distinction is unnecessary and generally unsatisfactory: firstly, it is not possible to define what makes a work qualify as 'literary', and secondly, a great deal of 'dialect writing' is of a completely different character, i.e. children's essays and radio scripts, as mentioned above in introducing the Shetland case. I shall therefore be using the term 'dialect writing' or 'writing in dialect' to cover the whole field. What is of the utmost significance, however, is obviously the writer's competence in and attitude to the dialect, as well as his/her *purpose* in using non-standard features in writing.

Most of the relevant research in this field has so far been devoted to works of fiction. Studies of fictional dialogue have, not surprisingly, been undertaken by literary as well as linguistic scholars. Literary critics have made penetrating analyses of what is known as 'point of view', looking at, for example, first vs. third person narrative, present or past tense, or past tense vs. perfect tense, and, to some extent, colloquial vs. formal style (cf. Traugott 1981). However, if *regional* features are discussed at all in literary criticism, they tend to be dealt with in a very sweeping fashion, such as the following characterization of Mrs. Gaskell's use of dialect: "Her men and women have characteristic modes of speech. Sometimes they are easy to recognize, as, for instance, by their dialect, which, *incidentally, she used well*" (A. Pollard, 1965, "Mrs. Gaskell", p. 254). What is even worse is the demonstration of general prejudice as to the inferiority of the characters speaking 'peasant language' in, for example, plays by John Synge (as heard in a lecture given at a department of English).

Although linguists have no doubt profited by the insights gained by literary studies of 'point of view', it should be obvious that a full understanding of the mechanism of dialogue, including non-standard features, can only be reached through *linguistic* analyses, representing various branches and levels of the discipline, such as pragmatics (of crucial importance for point-of-view studies), syntax, phonology/phonetics (including prosodic features), and dialectology/sociolinguistics. I want to emphasize at this point that what I am concerned with here is a purely *linguistic* study of *serious* attempts at giving a representation of non-standard language in writing. All in all I subscribe to the views taken and the methodology suggested by Sumner Ives in his classical study "A theory of literary dialect" (1950; somewhat revised for *A various language*, 1971).

The following is a presentation and discussion of his major points, with my own supplementary comments:

1. Serious writers employ dialect writing as a means of *realism*. They have tried to give an impression of literal accuracy, to show actual speech as actually used.
2. When representing a dialect, these authors have been acutely conscious that they are depicting something *peculiar*, something different from their own conception of the standard language. The characters who speak ‘dialect’ are set off, either socially or geographically, from the main body of those who speak the language. Usually the suggested difference carries some connotation of inferiority, but not always.²

“The use of non-standard dialect indexes a social status which is socially or educationally inferior, or it can function dynamically and signal familiarity and in-group solidarity” (Traugott 1981:1, discussing *tu/vous* variation). The latter function is apparent in Mrs. Gaskell’s “Mary Barton” (written in 1848, depicting a Lancashire setting), where *thou* is used with a certain degree of feeling and intimacy, *yo'* or *ye'* to friends of the same or of a lower social standing, and *you* is the most formal variant. It is obvious that Mrs. Gaskell did not use dialect to signal inferiority: in “Mary Barton”, for example, she makes a point of explaining lexical items in footnotes, sometimes by relating the dialect word to earlier uses by distinguished authors. That type of commentary is probably made with the specific and touching purpose of making the reader see the words, not “as signs of uncouth speakers”, but “by the dignity of the derivation from Lydgate, Chaucer or Skelton, for instance, she wishes to establish the dignity of those who use them” (cf. Melchers 1978:113). Consider also the use of non-standard language in a book like *Room at the top*, where John Braine makes the factory owner speak ‘broad’ Yorkshire, because it is part of his image — he is proud of having worked his way up.

Sometimes, as in *Sitting pretty* by Al Young, it is the standard speakers who are clearly odd, and indexed as ‘other’ by their pomposity (cf. Traugott 1981:5). This is clearly brought out in the famous dialect stories by the Swedish writer Gustaf Fröding, especially in “Dumt fôlk” (= “Stupid people”), written just before the turn of the century.

3. Nearly all examples of literary dialect are deliberately *incomplete*: the author is an artist, not a linguist or a sociologist. Dickens’s Cockney, for example, is “a literary stereotype which was seriously out-of-date at

the time he wrote and is now wholly obsolete" (Wells 1982:333). Dickens makes no attempt at representing some of the most salient features of Cockney. In representing Yorkshire speech, as in the character of John Browdie, he uses characteristics from various parts of the North in a rather unsystematic way. Not many will deny Dickens's greatness as a writer, but he was not considered worthy as a source of information for Wright's *English dialect dictionary*, where references are made to some writers of fiction, such as Mrs. Gaskell and Thomas Hardy.

4. Every variation from the conventional system of writing the language is a linguistic problem, and a valid theory of literary dialect must be based on linguistic evidence, especially that suggested by *dialect geography*. This is put forward by Ives as one of the great advantages of his time — earlier scholars in the U.S.A. had not had access to the regional patterns of American speech and the distribution of individual features, as brought out in the work on the linguistic atlas. As for Britain, the SED Basic Material was published in the 1960's, the Linguistic Atlas in 1978, and the Scots lexical atlas in 1975 and 1977 (the phonological atlas is due to be published later this year). Great strides have been made in linguistics since Ives wrote his article, and it now seems impossible to talk about dialectology without adding the social dimension. Methods developed in sociolinguistic studies can often be applied to the study of texts — cf., for example, variation in the second person pronoun as described above. In a study of Mrs. Gaskell's *Wives and daughters* I tried to apply network theory, social stratification, etc., and a recent study at my department describes linguistic variation in a work by Tolkien with the help of variable rules and implicational scales. Code-switching is perhaps the most obvious phenomenon to be investigated in written dialogue.
5. Ives argues that the starting-point should always be the *actual usage* in a text. After mapping this usage, generalizations can be made. Earlier scholars in the field seem to have worked the other way round. It seems to me that a combination of these two approaches should be the most useful — from our present knowledge of the dialects it is quite feasible to make hypotheses and predictions about an author's use of dialect features, also considering his background and prospective readers. As to generalizations, it should be realized that from the total lin-

guistic material the author selects typical features, sometimes even stereotypes, which can be compared to the sociolinguistic concept of *indexicals*. These features are often generalized by the author to such an extent that the literary dialect (in this case perhaps not dialect writing in general) is likely to be more regular in its variants than the actual speech which it represents.

6. The term *eye-dialect* seems useful for covering this whole field of study, i.e. the representation in ordinary writing of spoken dialect. It is indeed used by Wells (1982:529) in this way, but other scholars appear to use it in a more negative and limited way, indicating spellings that mean nothing at all phonetically — they are merely a sort of visual signal to the reader that the dialect speaker is not literate, as frequently done by 19th-century humorists. It may be that the connotation with eye-rhyme makes the term unfortunate in a more general sense, and I shall therefore avoid it.
7. The *limitations of conventional orthography* as a representation of phonetic value are obvious. This will be discussed in the following section.
8. As stated above, the importance of *the native speech of the author and that of his intended audience* is crucial for any interpretation.

It is difficult to find a better illustration of this point than that provided by Wells (1982:529), who exemplifies “the very different preoccupations of an American northerner and an Englishman in caricaturing a southern accent”, i.e. in the U.S.A., by the following quotations:

“Most of yew religious leaders ah haven’t seen in many yeuhs. Ah jus wus tawkin to Billi this aftuhnoon ... Ah, the greyaph heah tels the stawry... Faw the fust time in twelve yeuhs, ...” (L. Bruce, 1975: *The essential Lenny Bruce*).

“Ah, Apollo jars. Arcane standard, Hannah More. Armageddon pier staff”. (K. Amis, 1968: *I want it now*).

One obvious feature is, for example, the RP speaker’s use of *r* to signal a preceding long monophthong but no consonant, which would not make sense to a speaker with a rhotic accent. A close comparison of these two quotations will reveal that the foreign observer has provided a more detailed and ‘phonetic’ representation. ‘Pier staff’ for ‘pissed off’, for example, indicates the long, diphthongized character of the first vowel as well as the open character of the second, whereas it is dif-

ficult to see the point of Bruce's *aw* as in 'talk' and 'story', which seem to be examples of 'eye-dialect' in the restricted sense described above. However, Amis's text would be quite impenetrable to an American reader.

This brings us to the problem of assessing the reader's competence. A writer of English fiction presumably aims at a world-wide audience, and his/her dialect writing will then inevitably just look quaint to many readers.³ In the case of a place like Shetland, this is obviously less of a problem, in that most of the writing is clearly done for Shetland eyes. Nevertheless, most people find it difficult to 'read dialect', and it is obviously a skill worth teaching and practising as done in some Shetland schools. It would, in fact, be useful and interesting to make a more general study of this problem: how do different readers actually interpret the attempts by dialect writers to represent various accents?

9. The *value* of literary dialects to the student of language can be questioned. Obviously, nothing can replace a fieldworker's recordings, but writers are often good observers and have been found to provide additional information about many accents. As shown in the above examples, the most detailed and interesting representations are often given by outsiders. Mrs. Gaskell, for example, was able to give a truthful rendering (as compared with SED) of a dialect of which she had only a fortnight's experience (Whitby), and it is not surprising that she was used as a major informant for the *English dialect dictionary*. "The analysis of those literary dialects for which verifying evidence is available can quite possibly bring out clues to a more certain interpretation of evidence in historical developments" (Ives 1971:177).

Karl-Hampus Dahlstedt, the Swedish dialectologist and Professor of General Linguistics at Umeå, lists other values such as determining authorship, research in the emergence of the standard language, attitudes to correctness (as shown in the handling of dialect writing by some editors and publishers), 'explication de textes', the teaching of knowledge about dialects.

Using orthography to represent accents of speech — limitations and possibilities.

The solution to Mr. Stout's problem would undoubtedly be the use of

phonetic script — the completeness and elasticity of the IPA definitely caters to the Fair Isle sound system as well as any other. I have absolute faith in Mr. Stout's capability of mastering this craft and we have already given him a crash course in phonetics. However, with the exception of the faithful subscribers to *Le maître phonétique*, readers would probably not be very happy with this particular solution.

What are, then, the possibilities of featuring a non-standard accent in normal orthography?

As we know, there is no one-to-one correspondence between sounds and letters in the *standard* representations, but we learn to read and spell through acquiring knowledge of the conventions of each particular language. A native speaker/reader is usually able to tell the pronunciation of a word he/she has never come across. Thus a native English speaker would automatically read out the letter sequence 'bane' as /beɪn/, whereas a Swedish speaker would make it /*ba:ne/ (*denoting the characteristic 'accent 2' pattern of Swedish word intonation). It is beyond the scope of this paper to discuss the differences between spelling conventions in the standard languages of the world, but obviously these vary a great deal as to closeness to the phonetics/phonology of the language: in Finnish, for example, there is a very close connection, whereas the Faroese convention, also created at a very late stage, is of an 'etymological' character and very remote from the actual pronunciation.

As a starting-point for the further presentation and discussion, it must be emphasized that, unless a phonetic transcription is used, it is not possible to achieve perfect correspondence. The task is rather to devise a special, workable convention, not just an unsystematic, impressionistic use of quaint spellings. Let us look at some achievements and ideas in detail.

Rhythm and intonation

As demonstrated by the example from Kingsley Amis, it is actually possible to give the reader an idea of the *rhythm* of an accent. By writing 'standard' for the two words 'stand' and 'it', then placing a comma, and finally writing 'anymore' in two words, the first particularly accentuated by the inclusion of 'h', Amis manages to convey the characteristic, staccato rhythm of the south, with its combination of contractions and very long, stressed syllables.

As to intonation, it seems difficult to go beyond the use of full stops, question marks, exclamation marks and, possibly, hyphens (to indicate an

even, ‘flat’ pattern). In one of Shirley MacLaine’s autobiographical books she characterizes the speech of a man from Northern Ireland by putting question marks after every statement.

Another, more indirect way of featuring intonation is to make comments about it. This is done with great sophistication by the Orkney author George Mackay Brown:

“There is no longer a typical Orkney farmer or farmer’s wife; year by year they tend to look more like the town people. But you can still see an old countryman in Kirkwall on a Monday morning; he has red polished cheeks, he looks awkward in his best suit, he speaks in a slow rich wondering lilt, the sentence rising and breaking off at the crest: a good language for narrative and dialogue.” (*An Orkney tapestry*, p. 10)

The described pattern has been corroborated by instrumental analysis of Orkney speech.

Quantity

A characteristic trait of Swedish, Icelandic, and Norwegian phonology is that every stressed syllable has one and only one long segment, either a vowel as in Sw. ‘sil’ /si:l/ (= ‘sieve’) or a consonant as in ‘sill’ /sil:/ (= ‘herring’). A fairly general rule for stressed syllables gives (C)V:(C) or (C)V(C):. This is normally reflected in the spelling, so that long consonants are represented by two letters (the same two, with the exception of /k:/, which is represented by ‘ck’, at least in Swedish). The rules are slightly more complicated when final clusters are concerned. On the whole, it must be relatively easy to indicate differences from the standard — if a vowel in a dialect is relatively longer than the corresponding standard vowel, this can be shown by using only one letter symbol for the consonant. It is indeed exploited by dialect writers. The Swedish poet Gustaf Fröding, for example, has ‘liker’ for ‘licker’ (= ‘likes’), ‘litet’ for ‘littet’ (= ‘little’), and ‘hopas’ for ‘hoppas’ (= ‘hope’), the last two words being representations of ‘Swedish American’. All this obviously works for readers who speak Standard Swedish, not necessarily for everybody.

In the Standard English English pronunciation (RP), vowel length can be described as ‘free’ and there are no long consonants. (To some extent, however, this is oversimplifying the facts: in monosyllabic ‘feet’ of the structure (C)V(C) there is variation depending on whether the vowel is a so-called ‘long’ or ‘short’ one, and the length of the final C is affected.) Words like ‘pepper’ and ‘knitting’ should definitely not be pronounced with medial

C:, or worse: C: + C, the latter a true geminate. This is, in fact, one of the most salient characteristics of ‘Scandinavian English’, and may partly be ascribed to the influence of spelling. Obviously, the English convention of doubling letter symbols for consonants has partly the same function as the Scandinavian system — it indicates that the preceding vowel is relatively short.

For the recently published *Reader's digest great illustrated dictionary*, John Wells has devised an ingenious ‘respelling system’ for the pronunciation guide that uses for the most part familiar combinations of the alphabet. In this system, a consonant following a short vowel and followed by another vowel is always marked as double in the pronunciation guide: thus ‘river’ is shown as (rívvər), ‘batter’ as (báttər), but ‘give’ as (giv) and ‘bat’ as (bat). This obviously makes perfect sense to a reader who has an RP or near-RP accent, and probably works rather well for most English speakers. It is not quite clear to me, however, why ‘battle’ has to be (bátt'l). (') stands for ‘syllabic l’, but there seems to be no need for doubling the consonant symbol. Other instances are puzzling as well, such as (diss) for ‘dis-’ — would an English speaker read out (dis) differently?

Clearly, differences in consonantal ‘respellings’ cannot cater to all vowel-length distinctions in English. Letters representing vowel sounds are also doubled to indicate length, as in (laaf) for ‘laugh’.

An attempt has also been made at cross-dialectal adequacy. ‘Laugh’, for example, has the alternative pronunciation (laf) for Northern English.

Discussing quantity in his book *Regularized English*, Wijk is particularly concerned about the use of double consonants in *final position*, i.e. in the existing orthography. On the whole, it is only four or five consonant symbols that are doubled in this position, namely f, k, l, and s, possibly also z. While in the medial position consonant symbols are frequently doubled to indicate the short quantity of the preceding vowel, doubling in the final position cannot be said to have any function whatever in this respect (Wijk 1959:251). Consider, for example, ‘ebb’-‘web’, ‘back’-‘yak’, ‘bell’-‘expel’, ‘fuss’-‘bus’. However, Wijk does not suggest any major changes in the spelling conventions here, with the exception of words ending in -iff or -ll preceded by a or o, in which cases the vowel is mostly realized as long.

Looking at attempts by dialect writers to represent quantity characteristics, we find that consonant symbols are indeed often doubled to mark a preceding short vowel. For the Whitby pronunciations of ‘baby’, ‘water’, and ‘Newcastle’, Mrs. Gaskell has ‘babby’, ‘watter’, and ‘Newcassel’, and

Dickens has ‘lickly’ for ‘likely’ (Blackpool speaker) and ‘hoss’ for ‘horse’ (Suffolk).

I have left the Scottish system to the end of this section, since it is distinct from all others (except Ulster accents). There are no long—short oppositions of vowels as found in other English accents; on the other hand, duration contrasts (without *qualitative* differences) may signal morphological boundaries not found elsewhere. In fact, vowel duration tends to vary sharply according to phonetic environment. The general rule, known as ‘Aitken’s Law’, can be formulated as follows: “A vowel is phonetically short unless it is followed by ≠#, a voiced fricative, or /r/, in which case it is long (at least in a monosyllable).” (cf. Wells 1982:400). Long vowel duration is also retained if a morpheme-final vowel is followed by a suffixal /d/ as in *agreed*. Consequently, there is a phonetic distinction between *need* and *kneed*, and between *brood* and *brewed*.

None of this is shown in Wells’s ‘respelling’ system — what he is concerned with is not an overall accent distinction like the above, but only a lexical-incidental guide to pronunciation.

In 1947 a group of Scottish writers made a series of recommendations concerning the spelling of Scots which are embodied in *The Scots style sheet* and now generally employed by writers of ‘Lallans’, a revived form of Scots. Some features in the style sheet will be brought up in connection with the representation of vowels and consonants, but quantity is not highlighted in the recommendations. Obviously, a Scottish reader would automatically apply Aitken’s Law to the standard orthography. Not even ‘outsiders’ seem to represent Scots quantity patterns as compared with Standard English English, at least not to a great extent and then restricted to the occasional doubling of vowel letter symbols. Consider, however, J.B. Priestley’s description of a Scottish speaker:

“‘You’re going to stay here until that arm’s mended and you’ve had a nice rest and your nerves are quiet again.’ He still called them “nairrvés”. He still brought out those huge vowels and smashing consonants.” (*The good companions*)

Consonants

In John Wells’s respelling system, the idea is that a symbol stands for the normal, standard pronunciation of that letter. This goes for sequences of letters as well — thus (ch) stands for the affricate as in ‘church’. It follows that a word like ‘nature’ is represented as (náycher). Since (j) stands

for the initial and final sound in ‘judge’, (y) is taken to represent the approximant /j/ as in ‘yes’ (yess). As to the dental fricatives, the voiceless one is represented by (th), and the voiced by (th). The medial consonant in ‘vision’, ‘pleasure’, is represented as zh. For ‘exotic’ consonants from other parts of Britain, Wells give (hl) for Welsh ‘ll’, (kh) for Scottish ‘ch’ as in ‘loch’, but does not include a representation of the voiceless palatal fricative in Scottish, as in ‘light’. On the whole, this system seems to work very well, and should be adopted by any writer trying to convey features of pronunciation.

Clearly, it does not cater to the whole range of consonantal variation found in accents of English. There is, for example, no distinction made between clear and dark l — an allophonic difference in RP but also distributed differently over various accents of English. This distinction is never indicated by writers, nor do people in general seem to be aware of it.⁴ As to different qualities of /r/, frequent examples can be found of ‘failure’ to pronounce this consonant, making it into a bilabial approximant instead, such as ‘howwid’ for ‘horrid’. Dickens sometimes indicates a lingual roll by writing double r where it would not normally be found in the spelling, or ‘r-r’. It seems difficult to think of ways of representing other qualities of /r/.

Glottal stops are never shown. It is difficult to determine why Shaw and Dickens did not represent this characteristic feature in Cockney by some kind of symbol. One problem is, in fact, the dating of extended use of glottal stops in the dialect. It must be a fairly recent phenomenon, since it was not really exported to Australia. Actually, many speakers of English do not seem to be aware of this feature either. It would be easy enough to assign a written symbol to it. Barltrop & Wolveridge, in the popular book *The muvver tongue*, suggest (!) as in ‘ro!en’, ‘pota!o’. The ‘suspended’, glottalized /t/ as in realizations of the definite article in Yorkshire and Lancashire dialect, is normally represented as (t') as in ‘t'table’, and has always been one of the most salient features in dialect writing from these areas.

Vowels

Some noteworthy points in the Wells system: as in the case of consonants, spellings suggesting a certain vowel to the RP speaker are used as far as possible, i.e. (aw) stands for the vowel in ‘thought’, (ee) for the vowel in ‘meet’, and (eer) for the vowel in ‘fierce’; (i) denotes the short vowel in ‘grid’, whereas (i) stands for the diphthong in ‘price’, both according to the same established correspondence between spelling and pronunciation; (öö)

stands for the short vowel in ‘would’, and (öö) for the long vowel in ‘shoe’, whereas (o) represents the diphthong as in ‘goat’. Again, this system seems workable enough, even if it is not immediately transparent. The only symbol taken from outside the alphabet is (ə) as in the first vowel of ‘approve’. As to ‘exotic’ sounds Wells uses (ö) for anything like French ‘boeuf’ or German ‘schön’, and (ü) for anything like French ‘lune’ or German ‘Führer’.

The Scots Style Sheet recommends (ui) for the outcome of OSc /ø/, which is pronounced differently in Scotland, varying from a relatively close /i/ to /e/. Unlike Wells, the makers of the style sheet wish to use (ou) for /u:/. This is done to restore an old tradition, but is rather confusing, since it produces a convergence on StE forms in some cases, e.g. ‘out’. In this case the demands of historicity have taken precedence over those of autonomy, as pointed out by Caroline Macafee in her article “Nationalism and the Scots Renaissance now” (*English world-wide* 2:1, 1981:29-38).

As we know, not even the IPA system is quite adequate for representing all conceivable variations in vowel sounds, effected by slight movements of the tongue and lips, at least not without abundant use of diacritics. In Sweden, Denmark, and Norway traditional dialectologists have maintained that, owing to the enormous differences in vowel quality between the various dialects, it is useful to apply special symbols to special dialects. This is an easy way out, but a rather unfortunate one — if the data have not been recorded on tape, nobody will know the exact quality of the vowels any more.

Obviously, the limitations of normal orthography are even more striking. In general, however, it seems that dialect writers spend a lot more time devising spellings of vowel sounds than of consonants. From my own fieldwork in Shetland I also have the impression that dialect speakers themselves and lay observers of spoken dialect are much more aware of even subtle differences in vowel quality — such as a slightly fronted /a:/ versus a not fully back one — than they are of very striking consonantal differences, such as /kw/ instead of /hw/ for the beginning of the word ‘wheel’. Although H Dropping is a clear indexical for British speakers, not many other features involving consonants seem to be.

Looking at the attempts by some writers of dialect to represent characteristic vowels and diphthongs, we might note the following: Dickens’s Cockney lacks any indication of Diphthong Shift, which must have existed in the dialect in his time. This is indeed very surprising, since it is a relatively easy problem to deal with. Shaw, who is more complete here, writes

'tride' for 'trade' and 'civlawzytion' for 'civilization'. Wells (1982:334) criticizes him for the latter spelling, suggesting a too close beginning of the diphthong, but it is difficult to think of a better way of symbolizing it in ordinary letters. It is, of course, not necessary to turn it into a monophthong — here Shaw seems to be influenced by his own Anglo-Irish background.

As every EFL learner knows, there is a fairly close relationship between the spelling and pronunciation of English in the following way: every vowel symbol has its 'normal' short pronunciation in 'checked' monosyllables, such as 'hit', 'bed', 'pot', and its 'normal' long pronunciation in so-called 'open' syllables or before C + l/r, as in 'Peter', 'over', 'ruler', 'table', 'fibre'. This relationship can obviously be exploited by a dialect writer, as it has been in Wells's respelling system (cf., for example, Dickens's 'tride' above).

Obviously, there are also many combinations of vowel symbols that have a 'normal' pronunciation, such as 'ay', 'oy', which have been adopted by Wells and are used by dialect writers, at least those writing for an audience with a standard accent. When a writer goes beyond these normal correspondences, we can only try to interpret what the intentions are. When Dickens writes 'agean', 'deead' for a Yorkshire pronunciation of 'again', 'dead', a diphthong appears to be intended as in RP 'dear', but what does he mean by writing 'envys' instead of 'envies'? Possibly an /ai/ diphthong, but we can only make a guess here. If a diphthong was the intention, it would have been a better idea to add an 'e' at the end of the word.

It goes without saying that the greatest problems are created by sounds that have no equivalents in the standard at all, i.e. not just distributed differently over the lexicon. Such a vowel is, for example, a long, front, half-close to half-open one, as presumably meant by Emily Brontë in writing 'maister' for 'master'. Dickens's 'nae' for 'no' is simply an adoption of the Scottish convention.

Using 'umlaut' symbols such as (ö), (ø), or (ü) seems to be restricted to Shetland writing, and will be discussed below. It might be noted here that 'umlaut' apparently makes an extremely quaint and characteristic impression on English readers. In representing Scandinavian languages, umlaut is over-exploited to the same extent as Dickens exaggerated H Dropping. A Swedish beer is being launched on the American market as 'Nördik Wölfl'!

Addition; deletion; some aspects of connected speech

Features having to do with addition or deletion are obviously rather easy to handle — it is simply a question of adding or deleting a latter, as in ‘Hinglishmen’ for ‘Englishmen’ (Shaw) or ‘des’ for ‘desk’ (as in representations of Black English). The Scots Style Sheet strongly advocates the elimination of apostrophe spellings for deleted consonants, such as ‘de’il’, ‘lan’. These spellings are much disliked and criticised as unnecessary concessions to the non-Scottish reader. Caroline Macafee, in the above-mentioned article, also points out that they are concessions to the basic St E literacy of the Scots reader himself.

It is slightly trickier to represent the retention of a consonant in the actual pronunciation, which is extant in the standard spelling but not realized in the standard pronunciation. An example of this is the pronunciation of initial *kn-* as in ‘knife’, ‘knee’. In Shetland, as in some other areas in Scotland, a few old speakers still realize this as either a cluster /kn/ or use an epenthetic vowel /kən/. Dialect writers indicate this (if they do it at all) by ‘ken-’ for the epenthetic vowel and ‘k-n-’ for the cluster (also done by Wells for words like ‘knessel’). Metathesis, a salient feature in many Scottish dialects, including Shetland, is usually shown: ‘modren’, ‘pattren’ for ‘modern’, ‘pattern’. Except for the more or less ‘standardized’ representations of assimilation and weak forms, such as ‘gotta’, ‘yer’, and ‘I’ll’, these aspects of connected speech are seldom apparent in dialect writing, in the same way as they are not recognized by most lay observers of speech.

The case study: Writing in Fair Isle dialect Shetland phonology

“The vowel system of Scots comprises a subset of the vowel system of Scottish English ..., namely a basic /i, e, ɛ, a, ɪ, u, ɔ, ʌ/, to which certain local dialects add further items, giving anything from eight to twelve vowels in the system. To these must be added at least two, and commonly three, diphthongs.” (Wells 1982: 397)

It can be assumed that Shetlanders have, on the whole, a Scots system of vowels, though the actual sounds they make in particular words are often very different from vowel-sounds heard in Scotland. The Shetland vowel system has been particularly well described in stressed monosyllables closed by /t/ (J.C. Catford, J.Y. Mather, both working for LSS). In this position, as in fact in many others, Shetland dialects generally have a system of 12 vowels, 9 short and 3 long, plus at least two diphthongs (usually varieties of /ai/ and /oi/). The short vowels are: /i/ as in ‘beet’, /e/ as in ‘beat’, /ɛ/ as in

'met', /a/ as in 'fat', /ɒ/ as in 'pot', /o/ as in 'cut', /u/ as in 'about', /i/ as in 'bit', and /ø/ as in 'boot'; the long ones /ɛ:/ as in 'bait', /ɑ:/ as in 'fault', and /ɔ:/ as in 'boat'. Fair Isle is singled out by Mather as having an 11-vowel system, lacking the *fault/pot* contrast. Another characteristic of the Fair Isle vowel system is claimed to be a diphthongal realization of the 'bait' phoneme. It must be emphasized here that Aitken's Law is generally applicable, and that 'length' is only relative: the differences between vowels as presented above all relate to quality rather than quantity. On the whole, it would seem that Catford exaggerated the 'deviance' in Shetlandic syllable structure. For stressed monosyllables he claimed that the 'Scandinavian' type, as described above, was found, i.e. either VCC or VVC, and he accordingly writes his key words as 'fatt' for 'fat', but 'faat' for 'fault'. He does not discuss the possible existence of long consonants in medial position after a stressed vowel in words like 'pepper', 'letter' (cf. above, under 'Quantity'). In an earlier study (Melchers 1984) I have shown that there is no substantial evidence of 'Scandinavian' long consonants in medial position. There is, however, a curious distinction in many word pairs such as *paper-pepper*, *later-latter*. In Shetland dialect the vowels are often identical qualitatively, but the quantity shows a reverse picture as compared with RP, in that the stressed vowels in *pepper* and *letter* are longer (spectrograms show double length) than in *paper*, *later*. The consonant length seems to be identical. One LSS notation (notably from Fair Isle) shows *pepper* to have a long medial consonant, but there is no further evidence of this. As to the alleged Scandinavian character of stressed monosyllables, LSS notations as well as our own transcriptions, corroborated by instrumental analysis, do seem to indicate that in many cases the final C is more than twice as long as the vowel preceding it. It should be added that 'Auslautverhärtung' is a characteristic of Shetland phonology, and certainly contributes to the perception of final consonants as long or 'heavy'. A special feature in Fair Isle dialect is extreme focussing on the last word in a tone unit, realized either by extreme lengthening and diphthongization of the stressed vowel or by lengthening of the following consonant. This is brought out very clearly in Alec Stout's reading of his own text (cf. the first page of this article and the transcription below) and contributes to the popular perception of Fair Isle speech as characterized by extremely slow delivery.

One of the Shetland vowels, the /ø/ in the Mather/Catford system described above, is generally claimed to be 'Scandinavian' and people are very much aware of it. It occurs particularly in words of Scandinavian ori-

gin, but has been frequent in other words as well, such as the stressed vowel in ‘curious’. In Fair Isle this vowel is realized as half-close and could be transcribed by [y]. It is more open in certain positions, such as often before /r/, which is reminiscent of the ‘sinking’ effect of /r/ in most Swedish accents.

As to the consonant system, the following should be pointed out: like other varieties of Scots, Shetland dialect has retained the velar fricative /χ/, with a more palatal allophone [ç], when occurring after front vowels, as in ‘light’;

when /t/, /d/, /n/ are preceded by front vowels, they are described as ‘palatalized’ by LSS, but it could equally well be said that the vowels are diphthongized. These consonants are, in fact, *dental*;

/ʃ/ is mostly articulated somewhere in between the RP sound and Swedish /ç/ as in ‘tjära’. This has historical implications (cf. Hjaltland-Shetland and hjälpm (= helmet)-shalmet (= a white lamb with a black ‘cap’);

some clusters are retained, such as *kn-* ad *vr-*;

there are no dental fricatives — the initial consonant in ‘think’ is /t/ and ‘this’ has /d/. Some speakers are far from consistent here, or have developed their own lexical-incidental system, e.g. a dental plosive in ‘thing’, but a fricative in ‘anything’ (in word-final position after a vowel most speakers have fricatives);

there is no voiced affricate in words like ‘John’, ‘George’, but the voiceless counterpart as in RP ‘chop’ is used;

in some Shetland accents, particularly Whalsay and Skerries, /k/ and /g/ when followed by front vowels are realized as palatal fricatives;

initial *wh-* is generally /hw/, but in some accents /kw/.

Shetland spelling conventions

There are no firmly laid down principles as in the Scots Style Sheet. The first convention was really created in the making of the monumental etymological dictionary of the Scandinavian element in Shetland dialect by Jakob Jakobsen, a Danish-Faroese scholar. This dictionary is a remarkable achievement, with more than 10,000 entries and a wealth of idiomatic and ethnological material. However, neither the phonology nor the orthography is particularly convincing and very inconsistent. The spelling has a very strong Scandinavian bias, which is felt to be unfortunate by many Shetland writers today who would rather adopt a Scots standard with a few local modifications. Jakobsen uses ‘ø’, “da Norse O with the / through it” (as put

by the writer of a letter to the editor in a Shetland magazine, making an appeal for a Scots spelling), for all words containing the /ø/ phoneme, such as 'døless' (= 'indolent'), 'trøni' (= 'the snout of a pig'). He fails to indicate the retention of /k/ in initial *kn*-, and frequently doubles the consonant in spelling a word, as in 'plagg' (= 'garment'), without indicating a long consonant in the phonetic transcription and when there is no need for indicating a preceding short consonant. It is, in fact, quite difficult to find one's way in this dictionary. Many words are spelt in more than ten different ways, probably just reflecting idiosyncratic realizations by different informants, as in a word for snow-drift, taken down as 'murkavi'/'murakavi'/'murikav'/'murkava'/'murkafe'/'murkafa'/'murrakav'/'morrakav', etc.

However, Jakobsen's dictionary has been out of print for a long time, and very few Shetlanders have the chance of consulting it. The most commonly embraced convention now, also applied in the dialect crosswords, is that laid down in a more recent Shetland dictionary, viz. that by John Graham, published in 1979. Graham sets out his principles very clearly, as in the following points (pp. xxv-xxvii): spelling is at best a compromise; when there is no established convention, reading becomes laborious; the present state of private enterprise in spelling must place itself under some restraint; consistency is not necessarily a virtue, but without it spelling becomes an end in itself rather than a means to the real end which is communication. It follows, then, that

(1) An attempt should be made to base Shetland spelling on the convention most familiar to the reader, which for most Shetlanders will be English. (By this Graham seems to mean Standard *English* English — he is not an advocate of the Scots Style Sheet). An English word with a local pronunciation only slightly different need not vary from the normal English spelling. The reader will immediately identify it, and knowing that he is reading dialect will apply his own local pronunciation. It should be unnecessary, for example, to change 'calm' into 'kaam' or 'kaum', 'fixed' into 'fikst', or 'come' into 'kumm'. By stating this, however, Graham presupposes that the readers speak the same dialect as the writer — there is no way of showing local colour to outsiders if one adheres strictly to these principles.

(2) Outlandish forms, used perhaps in a misdirected zeal to emphasize the non-English character of the dialect, should be avoided as far as possible. Among these Graham include the unnecessary doubling of consonants, as criticized above. Yet Graham himself is not quite consistent here — why

write ‘groff’ and ‘trivvel’, for ‘example?

(3) Like the makers of the Scots Style Sheet he is not fond of apostrophes to indicate omissions, but approves of them if they show when a letter has been omitted from the normal *Shetland* usage.

(4) As to particular sounds, Graham prefers the symbol (ö) to (ü) for the /ø/ phoneme, since it is “a modified *o* sound”. However that may be, this preference seems sound enough, since internationally a similar sound is often represented by the former symbol in the standard convention. Graham sees no need for using the (ü) symbol at all, and reserves (y) for /j/ after consonants as in ‘gyo’ (= ‘creek’). However, he occasionally uses this symbol in other positions, as in ‘mylk’ (= ‘milk’), where it appears to be used for a diphthong rather like RP /ai/, but this usage is ambiguous.

Graham is a diligent writer of dialect himself. The following sentences from his writings will illustrate the general usefulness and relative freedom of his principles and will also show, as in ‘gjaan’, that he does not always practise what he preaches:

“While dis wis gjaan on, anidder wye o spaekin an writin wis shapin da local speech. Dis wis English — ösed by da Kirk, da laa-coorts an ida sköls.”

Most Shetlanders have a copy of Graham’s dictionary — in fact many of our informants seem to read up on some old words before we come to record them — and generally speak favourably of it. The objections, if any, are that it has ‘a Mainland bias’, i.e. that it is based on Graham’s own variety of Shetland dialect from an area not far from Lerwick, the ‘capital’ of Shetland, situated on Mainland, the largest by far of the inhabited islands. When asked to specify, informants usually give one or two examples of missing words, or words that have a slightly different meaning. Complaints about the spelling seem to be decreasing — it does appear as if this dictionary is becoming the convention. Many dialect writers consult it, but will often modify the spelling slightly according to the needs of their particular accent. One such writer is Isobel Holbourn from Foula, a small island off the west coast of Shetland Mainland. The following is taken from her prize-winning story about a naughty ewe: “I mind tedderin ane a dat yowes at hed lang yackles at we wir taen hame ta kill, an shü wis dat starty at shü tör da k-not trow da swill an cleared ootadaeks wi a piece ida tedder still trailin, ...”. In this accent, as on the West Side of Shetland Mainland, ‘she’ and a few other words are realized with a long, front [y:], phonetically quite distinct from the more open, central vowel in ‘tör’, and it makes sense to

indicate this by two different symbols. Notice also the indication of retained *k* as in *k-not*, a device not found in Graham's dictionary.

Other writers, such as the deservedly popular Rhoda Bulter, claim to spell quite impressionistically. Her writing certainly shows some inconsistencies and examples of 'eye-dialect' in the restricted sense, such as 'watter' for 'water', when a relatively long vowel is meant, 'waatch' as well as 'waander' standing for completely different vowels, and 'kyemps' and 'kyittums' when there is no trace of a glide after /k/ before a front vowel. It is difficult to know how this spelling works for a reader, since Rhoda Bulter is a well-known performer, often recites her poetry on Shetland Radio, and even has cassettes published together with her books. The same goes for the collection of dialect scripts that we have got from Shetland Radio. These, too, show many inconsistencies, especially relating to quantity. What is, for example, the point of writing 'Settirday' for 'Saturday', 'closs' for 'close', and 'rodds' for 'roads', when the quantity of the stressed vowel is very similar to the standard pronunciation anyway?

Clearly, every writer develops his/her own spelling habits, which are difficult to change. The most *spontaneous* material is found in children's writing, which is an extremely interesting source of information, especially when collated with their own reading of the texts. Here is a sample, written by a ten-year-old girl from Whalsay:

"I bide apun a peerie isle caad Whalsa
 Hits wan a Shetlan Isles
 Whalsa is a boanie place ta bide in
 Hits lent is aboot seevin miles
 ..."

'Wan' is a way of indicating a rounded vowel as in 'pot', and 'boanie' reflects the long, diphthongized vowel very well. Both these spellings were invented by the child. However, in Whalsay while attending an English lesson, I actually witnessed an instruction by the teacher on 'how to write Whalsa', at least as to some obvious features of the accent. This instruction sets out four conventions — no wonder the situation is confusing for a Shetland writer:

<i>English</i>	<i>Scots</i>	<i>Shetland</i>	<i>Whalsay</i>
knows	kens	keyns	tyeens
cake	cake	cake	tyœik
care	care	care	tyœr

game	gem	geym	dyemmm
going	gangin	gyaain	dyaaain

The ‘Shetland convention’ here is partly the teacher’s own creation — the spelling of ‘keyns’, for example, makes a point of ‘palatalized’ *n* as preceded by a front vowel (cf. above, under ‘Shetland phonology’), which is clearly unnecessary for a Shetland reader, who will interpret ‘kens’ in this way just as well. Again, this is just ‘eye-dialect’, trying to add some local colour to the writing. The same goes for the double consonant in ‘dyemmm’ there is no evidence whatever of consonants being longer in Whalsay than elsewhere. As to the local pronunciation of ‘cake’, ‘care’, these are rather good attempts at marking the main characteristics of Whalsay dialect, and there is clearly a need for going beyond Graham’s dictionary here.

Let us finally look at an attempt by a non-Shetlander to represent Shetland dialect in a text meant to be read by a wide audience. The following is taken from one of the transcripts in *Tocher*, a magazine published by the School of Scottish Studies, Edinburgh University:

“So he was comin hom one night wi his büddie o sillocks an waand, and as
he passed a certain knowe, he wis awaar ‘at they were a light sheenin oot
an he güd up tō examine this, an he saa ‘at...the trows wis dancin inside.”

Alan Bruford, who has transcribed most of the Shetland material, has informed me that his main policy is to convey the main features of the sound to someone accustomed to standard English without confusing him/her unduly. To indicate the [ɛ:] vowel he therefore avoids spellings like ‘main’, ‘mait’, and writes ‘maen’ and ‘maet’ — in fact, this is also the Shetland convention according to Graham, which Bruford does not refer to. He thinks of himself as catering in particular to an audience which is generally presumed to have some familiarity with mainland *Scots* forms, and stresses the fact that Shetland dialect is after all a variety of Scots. Therefore he obviously does not see the need for showing velar or palatal fricatives as in ‘night’, ‘light’ and deplores the influence of Jakobsen, as seen in some ‘writers’ use of double *k*, for example, which makes “Shetlandic into Scots written like Eskimo”. On the whole, Bruford is very sensitive about not using quaint spellings which might ‘degrade the speaker’. Therefore he shuns the use of *h* to mark a preceding long vowel, as in ‘Ah’, which “suggests slave-talk in the Uncle Remus convention”, which he finds can be said about ‘dis’ and ‘dat’ as well, although he uses the latter spelling convention when he actually hears it, but finds that the quality of the initial dentals is different

in the northern parts of Shetland. An apostrophe in 'at is unnecessary — the word is probably not a form of 'that' but related to Scandinavian 'at(t)'. Quite rightly, he observes the distinction between 'ö' and 'ü' and marks these characteristic features.

Obviously, the above represents dialect writing of a different kind, in that it is a representation of what someone has actually said rather than 'creative' writing, but it demonstrates the problem of the author's background and the competence of his readers. How would an Edinburgh or a London speaker actually read this text?

Writing in Fair Isle dialect

ðøniwai | 'opheli a: 'da:i ke:m ðə'run:d | en wi rɪgət wøs in wər sondi 'sθts
 en 'pø:lis t wər 'syn: | en 'styd 're:di fər wøli ε:rvins 'bos tə pik wøs 'ap ||
 dat mɔ:rnün də wɛ:dər wøzənə 'gyd. | en də sna: wəs 'ta:hwin | a:θɪŋ wəs
 θəm a slyfɪ 'mis: | en hɪt wəs "bra: 'ka:lɪ | də 'bos dju:li ke:m ðə'saret |
 en wi 'kle:m ε "bø:θrd |

The above is a phonetic transcription of the first lines of Alec Stout's writing, as quoted at the beginning of this article. Let us examine his orthographic representation in some detail, determine how successful he has been, and see where he stands in relation to the three conventions — English, Scots, and 'Shetland'.

It is immediately apparent that comparatively few words follow the standard English orthography. Those which do are mostly rather select words and those which are not part of an old, rural vocabulary, such as 'polished', 'suit', 'driver'. The orthography owes more to Graham's convention than to any other. There are some unnecessary apostrophes which would be disapproved of by Graham as well as the Scots Style Sheet advocates. The [ε:] vowel is represented by (æ) as in 'ready', 'Aervin', a spelling also found in Graham. However, the same sound is found in the word for 'weather', which is written as 'weddar'. A Shetland reader would probably read out the word as intended (cf. the account of the *paper-pepper* distinction above), but it is somewhat confusing. According to general Shetland practice, *th* is replaced by *d*, *t* whenever the realization is plosive rather than fricative, thus 'da', 'dat', but 'aa'thing', 'athin'. There is some discussion in Shetland as to whether the definite article should be given as 'de' or 'da'. The vowel usually seems to be realized as schwa, so this does not really matter. Like Graham, Alec Stout uses (aa) to indicate a relatively long [a:]. This works well and cannot be misinterpreted — as we saw, a

similar convention is used by Wells. The spelling ‘tehwan’ is puzzling — here (aa) could have been used to advantage. The reason for putting in *h* is either a reminiscence of Bruford’s ‘slave talk’ or a sophisticated observation of the slight preaspiration heard in the pronunciation of this word (cf. the phonetic transcription). Anyway, this spelling is unfortunate, since it clashes with the indication of a palatal fricative [ç] as in ‘assythte’. There is actually no need for a special symbol there either — as [X] and [ç] are found in complementary distribution, the same symbol, (ch), could be used. We would then get ‘asicht’. As to (ö) and (ü), there does not seem to be real need for distinguishing them in this accent, and they are used rather randomly in this text. One of the characteristics of Fair Isle dialect is marked in the spelling (uy) as in ‘duy’, ‘gruy’. This is well worth keeping, and should also be marked in many words where the standard has ‘air’ as in ‘hair’, ‘fair’. As to ‘Wullie’ for ‘Willie’, this is a correct observation of the extreme backness of the vowel before /l/, found in Fair Isle, Whalsay, and Skerries. There is no need to mark it for a Fair Isle reader, but it might be a good idea for a general Shetland audience. Quite a few words unnecessarily deviate from standard orthography, such as ‘kem’ for ‘came’. The long /a:/ varies in realization from back to front, but it is unnecessary to reflect this in the orthography. ‘Cauld’ should really be ‘caald’, but Alec Stout apparently perceives this back /a:/ as identical with the vowel in ‘folk’. This shows, anyway, that he is not a prisoner of the spelling convention.

On the whole, this writing proves to be rather successful, though obviously impressionistic. The above remarks are some indications as to how it could be made a little more systematic. Yet it is clear that competence in dialect reading can only be acquired through instruction — nobody could interpret a piece of writing like this automatically.

Some concluding remarks and general implications

In this article I have tried to point to some problems connected with the orthographic representation of non-standard speech. It must be emphasized, however, that this subject cannot be restricted to a phonological/orthographic discussion. Wherever non-standard writing is produced, it answers a purpose, and it is essential that a wide audience should be able to read it. Even if the intended audience is restricted to the actual speakers of a particular dialect, it takes some instruction to achieve competence in ‘reading dialect’, as it does in learning how to read any language — every time we learn a foreign language we have to re-learn reading up to point. It

is unrealistic to assume that this happens automatically — the ‘insight’ method should be advocated.

Several experiments have been carried out with the aim of teaching children to read their own dialect first. For various reasons, often emotional objections from parents, but just as often linguistic ignorance, most of these have failed (cf., for example, the touching account “Teaching Blacks to read against their will” (Stewart 1975)). In fact, most scholars seem to agree that using materials for teaching with a special spelling for Black English would not be good idea, at least not as to reflecting phonology.

Although there does not exist a thorough and systematic study of the question of the cross-dialectal adequacy of conventional English spelling, or of any other spelling, for that matter, it is obvious that nobody can help being exposed to and influenced by the standard orthography. It therefore seems natural to use it for teaching children to read. As some studies in Finland have shown, it is essential, however, that the children’s dialect is always considered in putting together teaching material, so that words are not used at an early stage which do not occur in the native dialect or have a form that would make the standard spelling opaque.

At a later stage, the reading of dialect could and should be introduced, with a solid linguistic instruction and application of other conventions. In Britain a movement has begun to impart more linguistic knowledge in schools (cf. the excellent series *Awareness of language* (CUP)). Knowledge of non-standard dialects, not just one’s own, and access to dialect writing should be part of this scheme.

NOTES

1) Traugott (1981:6) points out that many minority writers choose not to use BE at all: if it were used for third person narrative, a separate norm would be established, which would not help towards legitimizing Black culture as a full-blooded, fully visible but subcultural entity.

2) As is well known, dialect is often used to achieve a humorous effect. Even Robert Burns uses more standard language the more serious his poems are.

3) Deviant spellings are often seen as very disturbing as well: even Charlotte Brontë, in working over the 1850 edition of Emily’s *Wuthering Heights* modified and deleted some of her sister’s ‘Yorkshire orthography’ for the benefit of the readers.

4) Not even L Vocalization is indicated by Shaw or Dickens trying to represent Cockney, although this can be done quite easily, as in ‘miok’ for ‘milk’.

BIBLIOGRAPHY

- Desberg, P., Elliott, D.E., and Marsh, G. 1980. "American Black English and spelling." In: U. Frith (Ed.) *Cognitive processes in spelling*. London: Academic Press, 69-82.
- Graham, J.J. 1979. *The Shetland dictionary*. Stornoway: The Thule Press.
- Ives, S. 1971/1957. "A theory of literary dialect." In: J.W. Williamson and V.M. Burke (Eds.) *A various language*. New York: Holt, Rinehart and Winston, 145-177.
- Jakobsen, J. 1921. *Etymologisk ordbog over det norrøne sprog på Shetland*. København: Vilhelm Priors Kgl. Hofboghandel.
- Mather, J.Y. and Speitel, H.-H. 1975;1977. *The linguistic atlas of Scotland*. London: Croom Helm.
- Melchers, G. 1978. "Mrs. Gaskell and dialect." In: M. Rydén and L.A. Björk (Eds.) *Studies in English philology, linguistics and literature*. Stockholm studies in English XLVI. Stockholm: Almqvist & Wiksell, 112-124.
- Melchers, G. 1983. *NORN 1. A presentation of the project*. The Department of English, Stockholm University.
- Melchers, G. 1984. "Is the structure of the syllable in Shetland dialect 'Scandinavian'?" In: C.-C. Elert, I. Johansson and E. Strangert (Eds.) *Nordic prosody III*. Stockholm: Almqvist & Wiksell, 179-186.
- Orton, H. et al. (Ed.) 1962-71. *Survey of English dialects (SED)*. Introduction; Basic material (four volumes). Leeds: Arnold.
- Reader's digest great illustrated dictionary* 1984. London: The Reader's Digest Association Ltd.
- Stewart, W.A. 1975. "Teaching Blacks to read against their will." In: P.A. Luelsdorff (Ed.) *Linguistic perspectives on Black English*. Regensburger Arbeiten zur Anglistik und Amerikanistik. Band 9. Regensburg: Hans Carl, 107-123.
- Traugott, E. 1981. "The sociostylistics of minority dialect in literary prose". Paper delivered at Berkeley Linguistics Society Meeting, February 1981. Mimeograph.
- Wakelin, M. 1972. *English dialects: An introduction*. London: Athlone Press.
- Wells, J.C. 1982. *Accents of English 1-3*. Cambridge: CUP.
- Wijk, A. 1959. *Regularized English*. Stockholm studies in English VII. Stockholm: Almqvist & Wiksell.

THE REFLECTION OF LINGUISTIC STRUCTURE IN DUTCH SPELLING

Geert E. Booij

1. *Introduction*

Most linguists correctly assume that speech is the primary medium of language, and that writing is only a secondary, derived medium for conveying language.¹ This is already clear from the fact that every existing natural language is spoken, but not always written.² One might therefore be inclined to conclude that spelling is nothing but a graphical (written or printed) representation of speech, the audible form of language. It is the aim of this paper to show that this is a much too simple conception of spelling. I will argue that an orthographical system represents different levels of language structure, taking my examples mainly from Dutch.³

2. *Spelling and phonology*

Dutch spelling⁴ is often characterized as a more-or-less phonemic spelling, i.e. a spelling system in which letters or combinations thereof correspond to phonemes. By qualifying a spelling as phonemic, we already imply that written language not only abstracts away from the individual and situational variation in the pronunciation of words, but also from the effects of allophonic rules. That is, spelling does not represent speech, but a more abstract level of language structure.

However, it is not correct to consider Dutch spelling phonemic. For instance, the word *banaan* ‘banana’ can be pronounced as either [banán] or [bənán]. In the second phonetic form, the /a/ has been reduced to a schwa due to a rule of unstressed vowel reduction (cf. Booij 1977, 1981). A phonemic spelling would provide the orthographic form *benaan* (*e* stands for schwa in Dutch), since the schwa is a phoneme of Dutch. Yet, the correct

spelling is *banaan*. This may suggest that Dutch spelling represents underlying forms instead of phonemic representations, but this would be too hasty a conclusion. It appears that both underlying forms, phonetic forms and intermediate levels can be represented in Dutch spelling, as will be illustrated below. Therefore, it is much more adequate to characterize the Dutch spelling system by first establishing a typology of phonological rules, and then determining of which kinds of rule the effects are visible in spelling.

Rule typology is a neglected issue in standard generative phonology, but has become more important in recent developments in generative phonology such as Natural Generative Phonology (cf. Hooper 1976) and Lexical Phonology (cf. Kiparsky 1982). We may distinguish the following types of rules (cf. Booij 1981: ch. 8):

A. *Rules of sentence phonology* These are the rules which apply across word boundaries, such as the rules of external sandhi. Well known examples are the rules of voice assimilation in many Indo-European languages. Usually, these rules are optional.

B. *Rules of word phonology* These are the rules that only apply within words. The following subclasses have to be distinguished:

(i) *obligatory rules*, like:

- (1) Syllable-final devoicing
[– son] → [– voice] / ——)_o e.g. /hud/ ‘hat’ → [hut]
- (2) Devoicing of obstruents before tautosyllabic voiceless obstruents
[– son] → [– voice] / —— [– voice])_o e.g. /yud+s/ ‘good’ → [yuts]
- (3) Geminate simplification
 $C_i C_i \rightarrow C_i$ e.g. /et+t/ ‘eats’ → [et]

Rules (2) and (3) together derive the phonetic form of (*hij*) *biedt* ‘(he) offers’:

- | | | |
|-----|-----------------|---------|
| (4) | underlying form | /bid+t/ |
| | rule 2 | t |
| | rule 3 | Ø |
| | phonetic form | [bit] |

(ii) *morpholexical rules* (cf. Anderson 1974): rules which apply to a restricted (listed) class of words, or to words with a certain morphological property or specific morpheme. Well known examples are the German rules of Ablaut and Umlaut. Examples from Dutch are the rules which account for the allomorphy in the diminutive suffix, which has five allomorphs: *-tje*, *-je*, *-etje*, *-kje* and *-pje*:

(5)	traan 'tear'	traantje
	kam 'comb'	kammetje
	kant 'lace'	kantje
	koning 'king'	koninkje
	riem 'belt'	riempje

The specific form of the diminutive suffix is determined by the phonological form of the stem, but the rules only apply to diminutives. Morpholexical rules also tend to have (negative or positive) exceptions. For instance, several words have two diminutive forms (e.g. *weg* 'way' *-wegje/weggetje*, *kip* 'chicken' *-kipje/kippetje*), and in the pair *bloempje/bloemetje* (from *bloem* 'flower') we even find semantic differentiation: the meanings of the diminutive forms are 'little flower' and 'bunch of flowers', respectively.

(iii) *optional, style-dependent rules*. A relevant example is the rule of vowel reduction referred to above, which optionally reduces unstressed vowels to schwa.

(iv) *phonotactic rules*: rules which express restrictions with respect to the distribution of sound segments. For instance, in Dutch the nasal [n] never precedes a tautosyllabic /p/ or /k/: the nasal consonant is always [m] before /p/ and [ŋ] before /k/. Compare:

(6)	damp [dəmp]	*damk [dəmk]
	*danp [dənp]	*dank [dənk]
	*dangp [dəŋp]	dank [dəŋk]

Given the typology of rules outlined above, we are now able to give an adequate characterization of the Dutch orthographical system:

- (i) Dutch spelling abstracts away from the effect of the rules of sentence phonology, i.e. the spelling represents words as spoken in isolation.
- (ii) Dutch spelling does not represent the effect of optional, style-governed rules of word phonology, unless one wants to stress the informality of the kind of speech used (for instance in a written story).
- (iii) The effect of morpholexical rules is always visible in spelling. Con-

sequently, the diminutive suffix with its five allomorphs also has five different written forms. For instance, we do not write *raamtje* ‘little window’ (*/-tjə/* is generally assumed to be the underlying form of the diminutive suffix), but *raampje*, and *kannetje* ‘little can’, not *kantje*. This has a very desirable consequence: phonetic recoverability without recourse to syntactic context. For instance, if we were to spell *kantje* for *kannetje*, this form could also be interpreted as the diminutive form of *kant* ‘lace’. Such ambiguities can only be resolved by taking context into account. Similarly, *bloem* ‘flower’ has two diminutive forms, with different meanings, as pointed out above. If spelling abstracted away from the effect of morpholexical rules, both words would be spelled as *bloemtje*, a graphical form which would be ambiguous both with respect to phonetic and semantic interpretation.

(iv) Dutch spelling does not abstract away from the effect of phonotactic rules either. We do not write *damp* instead of *damp*, although the phonetic realization [damp] would be predictable. There is only one exception here: before a /k/ the velar nasal [ŋ] is spelled as *n*, e.g. *dank* instead of *dangk*.

(v) The last category of rules that needs discussion here is that of obligatory rules of word phonology. With respect to these rules the Dutch spelling system is rather inconsistent. For instance, we do abstract away from the effect of syllable-final devoicing of obstruents as far as /b,d,ɣ/ are concerned, but not with respect to the fricatives /v,z/:

(7)	rib ‘rib’	ribben (plural) <i>but</i> :	brief ‘letter’	brieven (plural)
	hoed ‘hat’	hoeden (“)	huis ‘house’	huizen (“)
	rug ‘back’	ruggen (“)		

Normally, Dutch spelling does not represent the effect of rule (2) (cf. *goeds* for /yud+s/ [yuts]), but in (*iets*) *vies* ‘(something) dirty’ the phonetic form [vis] is represented instead of the underlying form /viz+s/.

Finally, the effect of Geminate Simplification (3) is usually represented in spelling, as in *eet*, but not in words with the female suffix *-ster* /fits+stər/ ‘female cyclist’: *fietsster*.

Summarizing, Dutch orthography never represents the effect of optional rules (rules of sentence phonology and stylistic word-internal rules). It always represents the effects of morpholexical rules. The effect of obligatory word-internal rules is represented partially. Consequently, Dutch spelling may represent both underlying forms, phonetic forms and intermediate levels of phonological derivation.

An illustrative example of an orthographical form which represents an intermediate step of a phonological derivation is the spelling of the past

tense singular of verbs with stems ending in underlying /v/ or /z/, e.g. *roofde* ‘to rob’, past sg.:

(8)	underlying form	/rov+də/
	syllabification	(rov) _σ (də) _σ
	rule 1	f
	voice assimilation	v
	phonetic form	[rovdə]

In this case, the orthographical form *roofde* represents the intermediate level of derivation /rov+də/.

It will be clear from this description of Dutch orthography (i) that it represents different levels of abstraction and different levels of phonological structure, and (ii) that it is best described by making use of an independently motivated typology of phonological rules.

3. Spelling and hierarchical structure

So far, we described Dutch spelling as a system that represents linear sequences of sound segments at varying levels of abstraction. However, it also provides syntactic information. Capital letters mark the beginning of sentences and also mark proper names. Punctuation is used to mark the boundaries between constituents and the end of clauses. Spacing is used to mark word boundaries. In many cases these word boundaries indicated by spacing are both syntactic and phonological word boundaries, since usually syntactic word boundaries coincide with underlying phonological word boundaries. On the phonetic level, however, phonological word boundaries are often blurred by syllabification across word boundaries, as is illustrated by the phonetic syllabification pattern of *op een oor* ‘on an ear’: (ɔ)_σ (pə)_σ (no:or)_σ. When the number of syntactic word boundaries is not identical to that of phonological word boundaries, as is the case with compounds which contain more than one phonological word (see below), spacing gives priority to syntactic word boundaries.

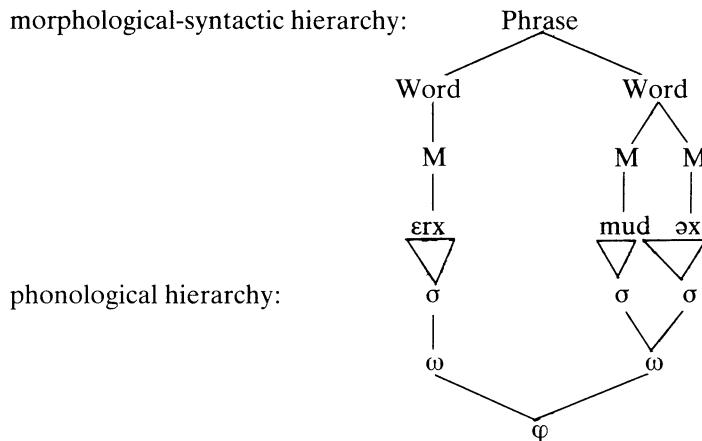
Thus, these orthographical conventions provide a wealth of information about the syntactic structure of a sentence. This is very important, because the phonetic form of a sentence does not always directly reflect its syntactic structure. Therefore, a graphic representation of the phonetic form of a sentence would be very inconvenient from the perceptual point of view.

One should realize, however, that sequences of sound segments are hierarchically structured in two different ways, since there are two hierarchies to be distinguished: a morphological-syntactic hierarchy on the one hand, and a phonological one on the other:

- (9) i the morphological-syntactic hierarchy: morpheme (M), word (W), phrase (Ph);
ii the phonological hierarchy: syllable (σ), phonological word (ω), phonological phrase (φ).

The idea that sequences of sound segments are structured in two different ways has been stressed repeatedly by Pike (e.g. Pike 1979), and has been also pointed out by, for instance, Halle and Clements (1982) and Booij (1983). Let me first illustrate the difference between these hierarchies by means of an example from Dutch, the structure of the phrase *erg moedig* 'very courageous'. The word *moedig* is derived from *moed* 'courage' by means of the suffixation of *-ig*:

- (10) morphological-syntactic hierarchy:



phonological hierarchy:

The difference between the two hierarchies is clear when we look at *moedig*: the internal morphological structure *moed-ig* is non-isomorphic with the phonological structure *moe-dig*.

Lack of isomorphism between the two hierarchies is also found in compounds and certain types of derived word. For instance, the compound $[[oog]_N [arts]_N]_N$ 'lit. eye doctor' is one syntactic word, but consists of two phonological words, *oog* and *arts*. If *oogarts* were one phonological word,

we would expect the syllabification pattern $(o)_\sigma (\gammaarts)_\sigma$, in accordance with the Maximal Onset Principle. The existence of a syllable boundary after *oog* is also proven by the phonetic form of this compound: $(ox)_\sigma (\alpharts)_\sigma$ with syllable-final devoicing of the underlying /γ/. It is known for many languages that the internal morphological boundary of compounds coincides with a syllable boundary. Since the phonological word is the domain of syllabification, we have to assume that compounds consist of more than one phonological word.

The official Dutch spelling marks syntactic boundaries only, and therefore there will be no space between the parts of a compound. On the other hand, many English compounds are written with an internal space, and thus English orthography gives priority to the phonological hierarchy. Interestingly, there is also a tendency among spellers of Dutch, in particular the less skilled ones, to separate the parts of compounds by spaces, in accordance with their phonological intuitions.

Certain classes of derived word also consist of more than one phonological word. Generally, one might distinguish two kinds of affixes, cohering affixes and non-cohering ones (cf. Booij 1983). Cohering affixes fuse phonologically with their stems, while non-cohering affixes form phonological words of their own. Certain Dutch suffixes, like *-achtig* '-like', *-heid* '-ness' and *-baar* '-able' are independent phonological words.⁵ Compare, for instance, the suffixes *-ig* and *-achtig*, which have the same meaning, with respect to syllabification (cf. Booij 1977: 1981):

- | | | | |
|------|-------------------------------------|-----------|---|
| (11) | $[[\text{rod}]_A \text{ig}]_A$ | 'reddish' | $(ro)_\sigma (dəx)_\sigma$ |
| | $[[\text{rood}]_A \text{achtig}]_A$ | 'reddish' | $(rot)_\sigma (\alpha x)_\sigma (təx)_\sigma$ |

Phonologically, words like *roodachtig* have the same structure as compounds, which is corroborated by the fact that they also have the same stress pattern: primary stress on the first phonological word, secondary stress on the second.

Certain prefixes show the same kind of behaviour with respect to syllabification, e.g. the prefixes *ont-* and *ver-*:⁶

- | | | | |
|------|----------------------------------|-----------------|--|
| (12) | $[\text{ont} [\text{aard}]_N]_V$ | 'to degenerate' | $(\text{ont})_\sigma (\text{a:rt})_\sigma$ |
| | $[\text{ver} [\text{as}]_N]_V$ | 'to cremate' | $(\text{ver})_\sigma (\text{as})_\sigma$ |

These observations with respect to the non-isomorphism between morphological-syntactic and phonological structure are not only relevant with respect to spacing, but also with respect to hyphenation rules. The hyphenation rules of Dutch crucially depend on the phonological hierarchy, i.e.

the hyphen comes after a syllable boundary, not after a morpheme boundary (for instance, we hyphenate *moedig* as *moe-dig*, not as *moed-ig*). Compare this to the English spelling which puts a hyphen after a morphological boundary, if possible (e.g. *morphem-ic*, not *morphe-mic*). A derived word like *roodachtig* will be hyphenated as *rood-achtig*. At first sight one might be inclined to conclude that here morphological structure is the determining factor. However, we have seen above that in this word a morphological boundary coincides with a syllable boundary, and hence we can conclude that, after all, the hyphenation rules of Dutch are consistently phonologically oriented. The same applies to prefixed words, which are hyphenated as follows: *ont-aard*, *ver-as* etc..

If possible, in Dutch orthography a higher phonological category takes precedence over a lower one. That is, Dutch spelling prefers a hyphen after a phonological word boundary to a hyphen after a syllable boundary. Thus, *rood-achtig* is preferred to *roodaach-tig*, and *ont-aarden* to *ontaar-den*.

In addition to sound sequences which form one syntactic word, but consist of more than one phonological word, we also find the inverse situation, sound sequences which are independent syntactic words, but are not independent phonological words. Clitics are examples of such sound sequences. For instance, in Dutch *-ie* is the weak form of the pronoun *hij* 'he'. This clitic always fuses phonologically with the preceding word, as is clear from its behaviour with respect to syllabification, e.g. *Kan-ie komen?* 'Can he come': (ka) _o (ni) _o (ko) _o (mən) _o. That is, *kan-ie* is one phonological word. Therefore, Dutch spelling requires the use of a hyphen before *ie*. Note also, that since *-ie* is phonologically dependent on a preceding word, it can never occur at the beginning of a sentence: **Ie komt* 'He comes' is impossible.

4. Conclusions

In conclusion, we have seen that Dutch orthography represents a number of levels of linguistic structure. First of all, it represents both phonological and syntactic information. The differences between the phonological and morpho-syntactic structurings of sound sequences appeared to be crucial for a proper account of spacing and hyphenation rules, and made it also possible to interpret the difference between Dutch and English spelling with respect to both spacing and hyphenation conventions, and to interpret the deviations from the spacing norm in the spelling

of compounds by unskilled spellers of Dutch.

Secondly, it appeared that rule typology rather than notions like ‘underlying form’ and ‘phonetic form’ is the key to an adequate description of the orthographical conventions of Dutch.

Generally, we have shown that orthography is much more than a representation of ‘how it sounds’. By means of abstraction from the effects of certain phonological rules in orthographical representation we achieve graphemic invariance for many morphemes of Dutch, which is undoubtedly convenient from a perceptual point of view, in particular because the skilled reader normally does not read with some phonological form as intermediary between graphemic form and meaning: skilled readers read like Chinese. However, this does not mean that graphemic invariance is an absolute condition for a workable spelling system. Above, we saw that Dutch morphemes ending in /v/ or /z/ have two graphical representations. Another source of graphical variation is the lack of a sufficient number of graphemes for the vowels of Dutch: long vowels are represented by the letters of their short counterparts in the following way: in closed syllables one vowel grapheme stands for a short vowel, two identical ones for its long counterpart. In open syllables, one vowel grapheme stands for a long vowel. Consequently, a morpheme like /bak/ ‘bake’ has two graphemic representations, *bak* and *bakk* (as in *bakken* ‘to bake’, infinitive). Similarly, /et/ ‘eat’ has the spelling forms *eet* and *et* (as in *eten* ‘to eat, inf.’).

Therefore, it is still an open issue which level of phonological abstraction would be the most optimal one for an orthographical system. That is, I have not claimed here that the present Dutch spelling system is the best one that one can think of. The moral to be drawn from the analysis of Dutch orthography given above is that those who think that a more phonetically oriented spelling would by definition be more adequate because such a spelling is the only feasible one from a scientific point of view, are simply wrong.

NOTES

1. See Tauli (1977) for criticism of the position that writing and speech are equivalent, independent media of language.

2. A possible exception is Modern Standard Arabic, which is nobody’s mother tongue and is not a spoken language at all (Coulmas 1983: 473).

3. See Smith (1980) for similar conclusions.

4. A general introductory book on Dutch spelling is Booij e.a. (1979).
5. Historically, these suffixes derive from words. Related suffixes are found in several Germanic languages, for instance German and Danish.
6. I do not want to claim that all Dutch prefixes are independent phonological words. Certain prefixes may be considered syllable-appendices to phonological words instead. The crucial point is that prefix boundaries always coincide with syllable boundaries.

BIBLIOGRAPHY

- Anderson, S.R. 1974. *The organization of phonology*. New York: Academic Press.
- Booij, G.E. 1977. *Dutch morphology. A study of word formation in generative grammar*. Dordrecht: Foris.
- Booij, G.E. 1981. *Generative Fonologie van het Nederlands*. Utrecht: Het Spectrum.
- Booij, G.E. 1983. "Principles and parameters in prosodic phonology," *Linguistics* 21, 249-80.
- Booij, G.E. e.a. 1979. *Spelling*. Groningen: Wolters-Noordhoff.
- Coulmas, F. 1983. "Linguistic problems of literacy-introduction," *Journal of pragmatics* 7, 467-77.
- Halle, M. and G.N. Clements. 1982. *Problem book in phonology*. Cambridge Mass.: MIT Press.
- Hooper, J.B. 1976. *An introduction to natural generative phonology*. New York: Academic Press.
- Kiparsky, P. 1982. "From cyclic phonology to lexical phonology." In: H. van der Hulst and N. Smith (Eds.) *The structure of phonological representations, Part 1*, 131-76. Dordrecht: Foris.
- Pike, K.L. 1979. "Universals and phonetic hierarchy." *Proceedings of the 9th Int. Congress of Phonetic Sciences Vol. II*, 48-52.
- Smith, P.T. 1980. "Linguistic information in spelling." In: U. Frith (Ed.) *Cognitive processes in spelling*. London: Academic Press, 33-50.
- Tauli, V. 1977. "Speech and spelling." In: J. Fishman (Ed.) *Advances in the creation and revision of spelling systems*. The Hague-Paris: Mouton, 17-36.

ZUR ALPHABETISCHEN ORTHOGRAPHIE ALS GEGENSTAND DER SPRACHWISSENSCHAFT

Herbert Penzl

English Summary

As the contributions to this volume indicate, orthography as a subject of linguistics implies: (1) its study as a system of visual reflexes of oral signs; (2) the creation of basic and auxiliary spelling systems (applied linguistics); (3) the study of the main evidence for historical phonology. As an example of this third type, the role and problems of orthography are outlined below for the history of the German standard language.

Runic orthography, e.g., in the horn of Gallehus (400 A.D.), and Wulfila's Gothic alphabet (4th century A.D.) are examples of the creation of orthographies in Old Germanic times from non-Germanic sources. The whole family tree of writing systems (Gelb 1962) in the western world is not linked to linguistic but cultural affinities. The orthographic systems in Old High German (750-1050) major texts like Tatian, Isidor, Notker are creations by highly competent native speakers proficient in Latin. The Middle High German (1050-1350) "normalized" spelling of our literary editions is a creation of 19th century German scholars.

Publications of the Early New High German period (1350-1720) show great regional variety in their orthography. Efforts of grammarians like Schottelius, Freyer, Gottsched and Adelung succeed in establishing a relatively uniform orthography by the early 18th century. Only few spelling changes follow until modern times. After 1945 vigorous attempts by some linguists can be observed in order to reform German spelling, particularly its general capitalization of nouns.

1. *Das graphische Zeichensystem.*

Es gibt wenige systematische Darstellungen von Problemen der Orthographie allein. Der Grund dafür mag sein, daß die meisten Probleme im wesentlichen als sprachspezifisch anzusehen sind. In diesem Bande ist eine repräsentative Auswahl von einschlägigen Studien vereinigt, die in ihrer Gesamtheit gut als allgemeine Einführung dienen könnten.

Orthographie oder Rechtschreibung ist nicht identisch mit Schrift, Schreibung oder Verschriftung, sondern bezieht sich stets auf ein Rechtschreibungssystem, eine Rechtschreibungs norm für Sprachen. Auch als solche ist Orthographie ein wichtiger Gegenstand für die Sprachwissenschaft, wobei wir dreierlei Hauptaufgaben oder Hauptthemen unterscheiden können. Das eine als das allgemeinste sieht in dem Schreibungssystem ein Zeichensystem, das ikonisch (abbildend) oder indexikalisch (anzeichnend) oder rein symbolisch sprachliche Elemente wiedergibt und zwar je nach Typus Laute, Lautgruppen, Suprasegmentales, Silben, Worte, Wortgruppen, Satztypen, auch semantische Begriffe. Unter den Schrifttypen der Welt finden wir Wortschrift, Silbenschrift, alphabetische Schrift, die oft teilweise kombiniert erscheinen. Wir wollen hier uns auf die alphabetische Schrift beschränken. Wir müssen aber im klaren sein, daß die griechische und dann die lateinische Schrift, Europas erste Alphabetschriften, aus semitischer (phönizischer) Silbenschrift und letzten Endes akrophonisch aus ägyptischer Wortschrift (Logogramme) und noch früherer Bilderschrift (Ideogramme) entstanden sind (vgl. Gelb 1962). Wir schreiben die Zahlen (1,2,3,4 usw.) als Wortschrift und die Namen der Buchstaben wie eine Silbenschrift (das B,C,D usw.) Moderne Kurzschriftsysteme haben in der Regel Zeichen für sehr häufige Morpheme (Silben oder Wörter).

Neben dieser mehr allgemein theoretischen Betrachtungsweise ist die der angewandten Sprachwissenschaft von besonderer Wichtigkeit in unserem Falle. Als Anwendung sprachlicher, besonders phonologischer Analyse ist die Schaffung von Orthographien, von Verschriftungssystemen für vorher nicht oder selten und unsystematisch geschriebene Sprachen und Dialekte, von Transkriptionssystemen (Kuhn 1981, S. 2) oder Hilfsorthographien, von Reformen bestehender Schreibungssysteme (Umschriftung) stets betrieben worden (2. unten).

Der dritte Type der Analyse der Orthographie (3. unten) ergibt sich aus den beiden anderen, dem theoretischen und pragmatischen. Das Hauptmaterial der historischen Phonologie der einzelnen Sprachen besteht

aus den historisch belegten Schreibungen, die nur vom Standpunkt von Schreibungssystemen und Schreibungsnormen zu verstehen sind. Von dieser Bedeutung der Orthographie soll hier vor allem mit Bezug auf die Lautgeschichte des Deutschen die Rede sein, wobei wir vom Germanischen ausgehen.

2. *Orthographierung und Graphierung.*

Für die alphabetische Schrift ist die Beziehung Schriftzeichen/Einzel-laut von überragender Bedeutung. Die Schaffung eines Schriftsystems, die sich auf der Analyse der Segmente des Lautsystems aufbaut, kann als ange-wandte Sprachwissenschaft bezeichnet werden und das Werk einzelner Per-sonen darstellen. Diese Art der Verschriftung ist einer unsegmentierten Bezeichnung überlegen. Für den Linguisten ergeben sich außer dem allge-mien graphemisch-phonemischen Verhältnis als Fragestellungen die Her-kunft und Auswahl der Schriftzeichen, deren Distribution und Variation, deren innere Verteilung und Systematik. Man hat parallel zu Phonem und Allophon auch Begriffe wie Graphem und Allograph oder Graph eingeführt (vgl. die Literatur bei Kuhn 1981, Fn. 5), was sich für die Be-bung als vorteilhaft erweist.

Bei den Kultursprachen Europas beginnt mit der Verschriftlung die di-rekt erfaßbare Sprachgeschichte. Immer wieder werden aber noch in der Gegenwart Volkssprachen ("Vernacularsprachen", Kuhn 1981) oder früher ungeschriebene Dialektvarianten von Linguisten verschriftet, alphabetische Schriftsysteme als Hilfsorthographien, Transliteration oder Transkription zu anderen Schriftsystemen verwendet. Als Beispiele für solche Volks-sprachen und Sprachvarianten denke man an das Grönländische, an amer-i-kanische Indianersprachen, afrikanische "Eingeborenensprachen", Balut-schi und Turkisprachen in Afghanistan, im Deutschen an das Schweizer-deutsche, Plattdeutsche, das Luxemburgische, Pennsylvaniadeutsche, im Englischen an den Shetlanddialekt usw. Die Faktoren der Verschriftlung sind selten rein linguistisch, besonders wenn es sich nicht um bloße Gra-phierung, sondern um Orthographierung, also um Normgraphierung han-delt. Hier wird unsere Fragestellung soziolinguistisch, oder sogar direkt sprachsoziologisch. Orthographie als Norm ist ja wie all Normen in der Ge-sellschaft auch soziologisch, also gesellschaftlich bedingt, wobei politische, kulturelle, religiöse, soziale, pragmatische, kommunikative Faktoren mit-wirken. Bildung, Änderung, Verbreitung, Variation einer sprachlichen

Norm sind aber Probleme der Soziolinguistik. In der althochdeutschen Periode des Hochdeutschen (5.1 unten) drangen fränkische Normen aus "politischen" Gründen auch ins Oberdeutsche, in der mittelhochdeutschen (5.2 unten) Periode aus kulturellen Gründen hochdeutsche Normen ins Niederdeutsche, im Frühneuhochdeutschen (6. unten) begünstigen zeitweise politische Faktoren, z.B. die kaiserliche Kanzlei, südliche Normen, kulturelle wie die "Literaturproduktion" und religiöse wie Luthers Reform mitteldeutsche Normen.

Die Diskussion über die "beste" Verschriftung ist deswegen nicht abgeschlossen, weil sie nur sprachspezifisch und pragmatisch entschieden werden kann. Theoretisch wäre eine Orthographie mit einer beiderseitig einmaligen ("eineindeutigen", engl. *biunique*) Beziehung Graphem/Phonem wissenschaftlich am besten vertretbar und für einen Muttersprachler am leichtesten erlernbar. Aber um diese Beziehung allgemein aufrecht zu erhalten, müßten den Varianten der Sprache auch Schriftvarianten der Orthographie entsprechen. Es kann die Orthographie sich entweder nur auf eine Sprachform beziehen oder durch unspezifische Zeichenwahl auch die interne Variation in der Aussprache oder bei Volkssprachen gar die Züge von Kontaktssprachen, wie z.B. die des Dänischen beim Grönäandischen, berücksichtigen. Eine Wortschrift kann, da unabhängig von spezifischer Lautung, von jedermann nach seiner Sprache gelesen werden; die chinesische Schrift ist so eine einiges Kulturband für die Sprecher der gegenseitig unverständlichen chinesischen Sprachen.

Lautwandlungen und Lautentlehnungen mancher Art erweitern im Laufe der Sprachgeschichte die Distanz zu der bereits üblichen Orthographie. Dialektlautungen, die nicht Grundlage der Orthographie bildeten, können später größeres Prestige und größere Verbreitungen erlangen. Es kann auch in der Sprachgeschichte zu neuen entscheidenden Sprachkontakten kommen, z.B. im Mittelenglischen nach 1066. In solchen Fällen ist ein Ersatz der alten Orthographie, eine Umorthographierung notwendig. thographierung notwendig.

3. *Lautgeschichte und Orthographie*

Für die historische Phonologie ist die Orthographie, d.h. das Schriftsystem mit seinen Zeichen das wichtigste Beweismaterial. Als Arbeitshypothese beruht das auf der Annahme, daß für den naiven Schreibenden der Vergangenheit als oberstes psychologisch-kognitives Prinzip ein direkter

Bezug Graphem/Phonem bestand, er also die distinktiven Einheiten seines Phonemsystems durch eigene Zeichen darstellen wollte.

Die Erfassung des Lautsystems der Texte gründet sich auf die Interpretation des Schriftzeichensystems der Texte: ohne diese ist keine phonologische Deutung möglich. Das ist im Anfang der schriftlichen Überlieferung noch einfacher als später. Also wurde in meiner Darstellung der Phonologie der althochdeutschen Dialekte (Penzl 1971) stets systematisch Schreibungssystem und zugrundeliegendes Phonemsystem getrennt behandelt. Systemlose Verschriftung, z.B. bei Namensformen oder Wortzitaten wie den langobardischen im lateinischen Text, macht die phonologische Deutungen besonders schwierig.

Abweichungen in der Schreibung zeigen die Verschiedenheit des Phonemsystems des Schreibers von dem phonologischen Grundsystem der verwendeten Orthographie. Die Schreibfehler von Schulkindern, also noch wenig geschulten, „naiven“ Schreibenden in der Gegenwart beweisen eindeutig die Beziehung Phonem/Schriftzeichen, bestätigen daher die Methode der historischen Phonologie für Texte der Vergangenheit. Für das Englishe vergleiche man dazu Hanna (1966), für das Deutsche z.B. das Material bei Zehetner (1977) mit seinen naiven Schreibungen bairischer Schulkinder. Die bairische Vokalisierung des [r] führte z.B. zu Schreibungen wie *Schandam* ‚Gendarm‘, *gefahn* ‚gefahren‘, *Nienbraten* ‚Nierenbraten‘, umgekehrtes *spätersiens* ‚spätestens‘ usw. (Zehetner 1977, S. 71-74).

Zeichenwandel und Phonemwandel sind nicht immer gleichzeitig und parallel. Wir können verschiedene graphische Typen des Zeichenwandels beobachten: z.B., Zeichenverschmelzung, Zeichenveränderung, Zeichenverbindung, diakritische Zeichenzusätze; Zeichenspaltung, Zeichenzusammenfall, Zeichenersatz, Zeichenschwund, umgekehrte Schreibung in der Zeichenverwendung beruhen als Typen bereits auf phonologischer Interpretation. Es gibt Zeichenwandel ohne Phonemwandel, z.B. im Sprachkontakt bei Übernahme von Zügen einer anderen Orthographie, z.B. im Altenglischen modifiziertes <d> und runisches <þ> für früheres <th>, im Mittelenglischen wiederum <th> für die altenglischen dentalen Spiranten und französisches <ou> für früheres <u> usw. Es gibt Phonemwandel ohne Zeichenwandel, z.B. mittelenglisch <i> und <ou> ([u:]) in *tide*, *house* zu den Diphthongen in der großen englischen Vokalverschiebung. Eine solche Zeichenumwertung von mittelhochdeutsch <i> <u> in *zit* ‚Zeit‘, *hus* ‚Haus‘ ist im Frühneuhochdeutschen (6.) nicht eingetreten: <ei> und <au> traten eine, weil dieser Zeichenwandel keine Verwirrung im Zei-

chensystem hervorbringen konnte; es gab keine große frühneuhochdeutsche Vokalverschiebung wie im Englischen.

4. *Altgermanische Orthographie*

Das altgermanische Runenalphabet wird vom größten Teil der Forschung auf ein norditalisches Alphabet des Lateinischen zurückgeführt. Viele der Zeichen in den Inschriften zeigen deutlich diese Herkunft, z.B. aus I R S T. Die Übernahme muß noch vor der Ausgliederung der germanischen Stämme erfolgt sein. Die Denkmäler zeigen deutlich ein Schriftsystem, eine Orthographie. Wir haben sogar einen Kurztext, dessen Sprache man als nordisch-westgermanische Ursprache ansieht (Penzl 1975, S. 69): die Inschrift des einen Goldhorns von Gallehus aus Nordschleswig (400 n. Chr.):

EC HLEWAXASTIZ / HOLTIJAZ / HORNA / TAWIDO

‘Ich Hlewagast, (Nachkomme des) Holt (aus Holt? der Holzschnitzer?) machte das Horn.’ (Penzl 1972, S. 64)

Die obige Transliteration interpretiert das entsprechende Runenzeichen noch als Sibilant <Z>, verschieden von <R>. Das Runenzeichen <X> deutet auf Reibelaut. E. Antonsen (1979) gab eine graphische Komponentenanalyse der Runenzeichen nach distinktiven Zügen.

Das Gotische von Wulfila's Bibelübersetzung des 4. Jahrhunderts ist in einem eigenen Alphabet geschrieben, dessen Erfindung ebenfalls Wulfila zugeschrieben wird. Die Ableitung der gotischen von griechischen Schriftzeichen ist ganz eindeutig. So ergibt sich die Frage, ob es nicht leicht möglich gewesen wäre, das Gotische mit dem griechischen Alphabet wiederzugeben und warum der dreisprachige Wulfila das nicht getan hat. Es besteht kein Beweis, daß Wulfila Goten (“*Gothi minores*”) bereits eine runische Schrifttradition besaßen. Der Grund war wohl, daß Wulfila in seiner kleinen Sprachinsel im fremdsprachigen Raum ein eigenes Schriftsystem zu seiner Übersetzung schuf, um auch durch seine Verschriftlung die Nationalität seiner Gemeinde gegenüber dem kulturellen Übergewicht des Griechischen behaupten zu können. Die Übernahme der Wulfilaspalte und Wulfilaorthographie durch die Ostgoten Italiens zeigt dasselbe Bestreben gegenüber dem Lateinischen. Gotisch wurde sogar in Urkunden des 6. Jahrhunderts neben dem Lateinischen verwendet: bei keinem anderen germanischen Stamm findet sich solche Sprachloyalität.

5. *Althochdeutsche und mittelhochdeutsche Orthographie*

5.1. *Althochdeutsche Verschriftung*

Die überlieferten althochdeutschen Texte (750-1050) zeigen uns die Methoden der Verschriftung. Da andere Quellen, etwa zeitgenössische grammatische (orthoëpische) Beschreibung, Metrisches und Reime, Entlehnungsentsprechungen überaus spärlich sind, ist das Beweismaterial der ahd. Phonologie (Penzl 1971, S. 19ff.) fast ausschließlich die Orthographie der Texte. Alle phonologischen Kontroversen der Forschung betreffen im Althochdeutschen, aber auch im Altenglischen, Altsächsischen, fast gänzlich die Deutung des orthographischen Materials. Nach der mittelalterlichen Ansicht weist jeder alphabetische Buchstabe *figura* (Form) und *potestas* (Lautwert) auf.

So zeigten sich m.E. einige Forscher der Gegenwart als ausgesprochene "Figuralisten" (Penzl 1982): sie nahmen z.B. in der Frage das ahd. *i*-Umlauts die Stellung ein: kein Schriftzeichen — kein Umlautphonem.

Die Forschung untersuchte alles Einschlägige der Verschriftung: orthographische Zeichenwahl, Zeichenverteilung, Zeichensystem, Zeichenvariation, Zeichenwandel. Was die Zeichenwahl anbetrifft, erwiesen sich die einfachen Zeichen des lateinischen Alphabets nicht ausreichend für die ahd. Phoneme, so daß vorhandene (*ph th ch*) und neue Zeichenverbindungen (*ee ie ia uo hh uu zz*) oder einfache Zeichen mit mehrfachen Lautwerten, z.B. <a> als /a/, /a:/, verwendet wurden. Ein Text des 11. Jahrhunderts, die sogenannten "Altdeutschen (Pariser) Gespräche" enthält in altfranzösischer Orthographie recht fehlerhafte ahd. Dialoge und Wendungen zum Gebrauch französischer Reisender (Braune 1969, S. 9). Die Übereinstimmung der französischen und ahd. Zeichenwahl für die beiden Zischlauten /s/ /Z/, z.B. *luzzil* (mhd. *lützel*, engl. *little*), *guez* 'weiß', *ros* 'Roß', *taz* 'das', lässt in diesem Falle altfranzösischen Einfluß bei der allgemeinen ahd. Zeichenwahl vermuten. Die Methode des französischen Verfassers entspricht übrigens der amerikanischer Linguisten im 2. Weltkrieg, die in Sprachführern ("Language Guides") für militärische Benutzer eine Umschrift in englischer Orthographie verwendeten, z.B. für "Wie geht's" VEEGATes usw.

Die Untersuchung der orthographischen Zeichensysteme in ahd. Großtexten wie dem im wesentlichen ostfränkischen Tatiantext, in der alemannischen Benediktinerregel, dem fränkischen Isidor, dem südrheinfränkischen Otfrid, dem spätemannischen Notker erweist sie als sorgfältig ge-

plante, teilweise bewunderswerte Leistung in der Verschriftung der ahd. Dialekte (Penzl 1971). Die Planung zeigen auch deutlich Otfrids Schreiben an Erzbischof Liutbert und Notkers an den Bischof Hugo von Sitten. Bei kleineren Denkmälern sind ähnliche Verschriftungstendenzen zu bemerken, aber schwieriger zu erkennen. Bemerkenswert ist z.B. in der Benediktineregel die Wiedergabe von Langvokalen durch Doppelschreibung (*aa ee ii*) auch in schwächer akzentuierten Silben, bei Notker die systematische Unterscheidung von Langvokal durch Akzente (â/á; ê/é usw.) und sein berühmtes Anlautgesetz mit *p/b*, *t/d*, *k/g*/, auch *f/v* Variation je nach Endkonsonant des vorhergehenden Wortes, was allophonisch (Penzl 1971, S. 103f.) und phonemisch (Moulton 1979) gedeutet wurde.

Der spätalthochdeutsche und frühmittelhochdeutsche Zusammenfall der Kurzvokale in schwachakzentuierten Silben zeigt sich durch Schreibungs zusammenfall, durch umgekehrte Schreibung (engl. *inverse spelling*) und zahlreiche Gleichschreibungsfälle (Dittographien), z.B. im Merigarto des 11. Jahrhunderts (Braune 1969, S. 140-142, Voorwinden 1973): *sagan* ‘sagen’, *wisin* ‘Wiese’ (Akk. Sing.), *vili* ‘viel’, *vuinis* ‘Weins’ (Gen Sing.). Die Endung der 3. Person Pl. Prät. Indik. (ahd. *-un*) wird <-an> (*sprungan*), <-en> (*durchrunnen*), <-in> (*trogin*), <-un> (*runnun*, *skinun*) geschrieben.

5.2. Mittelhochdeutsche Schreibungen.

Die Orthographie der mittelhochdeutschen Handschriften (1100-1350) unterscheidet nicht Langvokal und Kurzvokal, selten die schon ahd. entstandenen Umlautvokale durch eigene Zeichen. Aber in den neuhighochdeutschen Textausgaben mhd. Epen und Lyrik finden wir eine von Germanisten des 19. Jahrhunderts geschaffene Normalorthographie, die systematisch zwischen *i/i*, *û/u*, *ê/e*, *ô/o*, *â/a* unterscheidet und alle Umlautvokale durch Zeichen wiedergibt, z.B. den Umlaut von *û* als <*iu*> (*hûs*, Pl. *hiuser* ‘Häuser’, z.T. schon bei Notker), von *u* als <*ü*> usw. Hier haben wir also einen Fall von normalisierender Neuverschriftung, eine Art Transliteration. Ähnliches hat man für althochdeutsche Texte nie unternommen, nur einen ganz schlecht überlieferten Text wie das alemannische Georgslied hat man immer umgeschrieben (Braune 1969, S. 132). Die meisten mhd. Handschriften sind oberdeutsch und die Texte zeigen ein einheitliches Zeichensystem, das wenig Schlüsse auf dialektische Lautung zuläßt. Doch finden wir auch mitteldeutsche Texte mit abweichender Orthographie, z.B. das riuarische “Die Lilie” mit <*i*> als Längebezeichnung in *gain* (normalmhd.

gân) ‘gehen’, *stain* (mhd. *stân*) ‘stehen’, *doit* neben *dot* (mhd. *tôd*, *tôt*) ‘Tod; tot’, *loin* (mhd. *lôn*) ‘Lohn’.

6. *Frühneuhochdeutsche Orthographie.*

Die frühneuhochdeutsche Periode (1350-1720) lässt sich als die Periode landschaftlicher Schriftstile charakterisieren (Penzl 1984). Ihr Ende am Anfang des 18. Jahrhunderts ist der Beginn einer hochdeutschen Schriftsprache mit im allgemeinen einheitlicher Orthographie ohne deutliche landschaftliche Variation. Die Einigung der hochdeutschen Aussprache ist allerdings nicht vor das Ende des 19. Jahrhunderts anzusetzen (Penzl 1975, S. 131f.)

Die verhältnismäßige Einheitlichkeit am Ende der Periode ist dem aktiven Eingreifen der Fachleute zu verdanken. Für die Lese- und Schreibmeister der Zeit, die beruflich gegen das noch überwiegende Analphabetentum der Erwachsenen ankämpften, mußte eine Vereinheitlichung der Orthographie mit nahem Bezug von Lautung und Schreibung überaus wichtig erscheinen; man vergleiche dazu die Texte bei Johannes Müller (1882). Die grammatischen Beschreibungen und orthographischen Empfehlungen bei J.G. Schottelius 1641, H. Freyer 1722, J.Ch. Gottsched 1748ff., J.Ch. Adelung 1782 sind weitere Stufen auf dem Wege zur hochdeutschen Einheitsorthographie. Es fehlt aber keineswegs an abweichenden Vorschlägen, schon das 16. Jahrhundert kannte “gelehrte Orthographiereformer” (Jellinek 1913, S. 56ff.)

An allgemein orthographischen Zügen der Periode sind zu erwähnen: die zunehmende regelmäßige Verwendung von Satzzeichen wie Punkt, Beistrich (Komma), Doppelpunkt, Strichpunkt, Fragezeichen, Ausrufezeichen. Die nützliche Virgel, vor allem ein Trennzeichen für Wortgruppen überlebte die fröhnhd. Periode nicht. Die Satzzeichen geben die Satzintonation an (Penzl 1984a). Die Herausgeber haben in altdeutschen Texten vielfach als Lesehilfe neuhochdeutsche Satzzeichen eingeführt, dadurch aber oft eine einseitige sprachliche Interpretation durchgesetzt, die nur didaktisch, aber nicht wissenschaftlich vertretbar ist.

Typisch für die Periode sind *<v>* *<j>* als graphische Anlautsvarianten von *<u>* *<i>*, z.B. *vff* ‘auf’, *jn* ‘in’, und vielfache “Konsonantenhäufung”, z.B. *<dt>*, noch jetzt in *Stadt*, *Städte*, *Verwandter*. In vielen Texten werden noch mhd. *ei* (*ein*, *stein*) und mhd. *î* (mhd. *zît*) als *<ai>* (*stain*), bzw. *<ei>* (*zeit*) unterschieden usw. Umlautsbezeichnung, Längeschrei-

bung der Vokale, auch die Schreibung der Sibilanten (Zischlaute, mhd. *s*, *z*, *tz*) zeigen mannigfache Unterschiede, sogar in Offizinen derselben Stadt.

Die Großschreibung aller Substantive entwickelte sich im Laufe des 16. Jahrhunderts aus der Schreibung am Satzanfang, von Gott, Eigennamen, Texttiteln u. dgl. (Penzl 1984, §26, S. 37). J. Becherer war im Jahre 1596 der erste Grammatiker, der sie verlangte; Stephan Ritter folgte ihm 1616 damit in seiner *Grammatica Germanica Nova*.

7. Pläne einer Rechtschreibreform.

Die Orthographie des Deutschen war schon in ahd. Zeit die Schöpfung von Fachleuten (5.1 oben). Das gleiche gilt in verstärkten Ausmaß für die frühneuhochdeutsche Periode (6.). In Einzelheiten setzte sich im 19. Jahrhundert diese Tradition fort. Jacob Grimm kämpfte für das historisch-etymologische Prinzip in der Rechtschreibung, z.B. in der Längebezeichnung. Er wollte <ie> in *Liebe*, *ziehen* (mhd. *ie*), aber nicht etwa in *sieben*, *Riese*, *Sie* (mhd. *i*), <h> in *Stahl*, *zehn* (mhd. *h*), aber nicht in *nehmen*, *stehlen*, *nähren*. Dagegen traten Germanisten wie R. von Raumer, W. Wilmanns und ein Praktiker des Lehrfaches wie Konrad Duden für das phonetische Prinzip, also für eine womöglich einheitliche Bezeichnung für die gleichen Laute ein. Erst die Rechtschreibtagung von 1901, die sogenannte 2. Berliner Orthographiekonferenz (die erste war schon 1876), beseitigte <th> statt <t> in deutschen Wörtern wie *Thal*, *Theil*, *Thon*, *Thür*, wo es sich vor Langvokal oder Diphthong erhalten hatte.

Es ist also durchaus im Rahmen einer alten Tradition, wenn auch in der Gegenwart Sprachwissenschaftler eine neue Orthographie, ein neues Schriftbild schaffen wollen. Hier haben wir zuerst als Resultat nach zahlreichen Tagungen (vgl. Nerius 1975, S. 71ff.) die sogenannten Stuttgarter Empfehlungen von 1954, dann die Wiesbadener Empfehlungen von 1958 und zahlreiche Veröffentlichungen von Fachleuten wie G. Augst, W. Mentrup, D. Nerius, Maria Hornung, Eugen Wüster. Im *Jahrbuch für Internationale Germanistik* (vgl. Band, 6: 1 (1975), 6: 2 (1975), 9: 2 (1977) ist diese Frage oft behandelt worden (Penzl 1985).

Ist nun die deutsche Orthographie wirklich im allgemeinen so reformbedürftig? Als Maßstab legt man, trotz Kritik der Generativisten, stets die Beziehung Aussprache (Lautung) und Schreibung an. Die deutsche Rechtschreibung, wie oben ausgeführt, das Resultat einer Jahrhunderte langen Entwicklung, schneidet bei einer Analyse gar nicht schlecht ab. Es gibt viel-

leicht Schönheitsfehler: die Längebezeichnung der Vokale ist inkonsistent (*Ahle, Aal, aber, Bart, Maße* gegenüber *Masse; sie, wir, Benzin*). Die Fremdwörter haben im allgemeinen ihre ursprüngliche Rechtschreibung behalten. Zum Unterschied vom Mittelhochdeutschen (5.2 oben) werden Morpheme einheitlich (“morphophonemisch”) geschrieben: *Berg* [k], *Berges* usw., mhd. *berc, berges*. <chs> steht für [ks] in Wörtern wie *Ochse, Dachs, Fuchs*. <v> für /f/ z.B. in *viel, ver-, Vater*. <ig> steht statt <ich> in *richtig* usw. Die Schreibungen *Lerche/Lärche, Felle/Fälle* und *Leib/Laib, Seite/Saite* unterscheiden Morpheme bei gleicher Lautung. Wir können feststellen, daß die deutsche Rechtschreibung im allgemeinen durchaus “benutzerfreundlich” und gut ist. Die Reformfreunde der Gegenwart denken also kaum daran, die erwähnten Schönheitsfehler, wenn es solche sind, zu beheben. Einige Fremdwörter scheinen ihnen in der Schreibung verdeutschbar; wir haben ja bereits *Fotograf, Sinfonie, Frisör* in der Rechtschreibbibel, dem “Duden” (vgl. auch Drosdowski 1975). In der Schweiz ist unter französischen Einfluß das Zeichen <ß> allgemein durch <ss> ersetzt worden (vgl. Kuhn 1981, s. 13, 2.2 (ii) zu K. Pikes ‘social’ goals’).

Der Hauptangriff der modernen Linguisten richtet sich gegen die Großschreibung aller Substantive: die Einführung der “gemäßigten” (?) Kleinschreibung soll sie auf Eigennamen beschränken. Durch die Aufgabe der Großschreibung ginge natürlich keine phonologische Unterscheidung verloren, aber nicht nur eine rein “figuralistische”. Es ist kein Zufall, daß gerade im Deutschen schon seit dem 16. Jahrhundert die Großschreibregel für Appellative erhalten geblieben ist. Für die deutsche Wortbildung ist die Substantivierung als Wortartwechsel besonders wichtig und die Großschreibung ein willkommenes Signal: *Reisender, Gefangener, Leben, Junge, Übel* usw. Für die Klammerkonstruktion deutscher Wortgruppen mit dem Substantiv als Zentrum ist die Großschreibung besonders für fremdsprachige Leser eine viel beachtete Hilfe: z.B. “diese von allen stets berücksichtigte Theorie”. Der Schweizer Germanist R. Hotzenköcherle (1955) hat auf diese Tatsache überzeugend hingewiesen.

Warum auf einmal diese heftigen Angriffe auf die Großschreibung? Weil sie schon J. Grimm als pedantisch empfand? Weil sie eine Erschwerung des Schreibenlernens, wenn auch nicht des Lesenlernens, bedeutet? Weil die Unterscheidung Eigename/ Appellativ (engl. z.B. *this University*, d.h. in meinem Fall *University of California, Berkeley* und *this university*) sehr wichtig ist? Weil die Kleinschreibung das Schriftbild verschönert (Figuralismus)? Der wahre Grund für die orthographischen Reformbestre-

bungen im Deutschen muß wohl sprachsoziologisch sein. Sprachloyalität im Sinne J. Fishmans (vgl. Fishman 1977) ist auch ein Faktor in Schriftbild und Orthographie.

Vergleichen wir die Lage im Englischen. Noch nie hat das Streben nach einer "simplified spelling" so wenig Anhänger gehabt wie in der Gegenwart. Unter den Fachleuten sind mir keine bekannt. Das kann nicht auf Chomsky-Halles Rühmen der englischen Orthographie als idealer, d.h. wohl hier morphophonemischer Wiedergabe der "Tiefenstruktur", z.B. gleiches <crim> in *crime*, *criminal* usw., zurückgehen (Chomsky & Halle 1968, S. 49). Die Inkonsistenz der englischen Rechtschreibung zeigen immer neue Vorschläge zur Hilfsverschriftung und Transkription. Sogar ein eigenes "Initial Teaching Alphabet" ist in England im Elementarunterricht erprobt worden (Downing 1967): die Ergebnisse beweisen eindeutig, wie eine Reform der englischen Orthographie ihre Erlernbarkeit sehr verbessern könnte (Kuhn 1981, S. 82-85). Und doch denkt jetzt kaum ein Laie oder Fachmann in angelsächsischen Ländern mehr an Reform.

Was das Deutsche anbetrifft, hat, wie z.B. Keller (1979) hervorgehoben hat, keine Sprache in der Welt seit dem letzten Weltkrieg mehr an Sprechern, Verbreitung und Ansehen verloren. Der Drang scheint über groß, fremde Sprachmuster, besonders das englische, nicht nur in Wortschatz, Syntax, Idiomen, sogar in der Orthographie siegen zu lassen, vor allem, wenn es sich um eine nur deutsche Eigenheit handelt. Die Großschreibung der Substantive gilt als solche; das hat deutlich 1948 die öffentliche Diskussion über die Abschaffung der Großschreibung in Dänemark bewiesen.

8. Schlußbemerkungen: *figura* und *potestas*.

Orthographie als Zeichensystem mit Form (*figura*) und Lautwert (*potestas*) seiner alphabetischen Glieder ist stets ein Gegenstand der Sprachwissenschaft gewesen. Die Verschriftung oder Umschriftung gesprochener Sprachen und Dialekte, die oft auf genauer Analyse von Phonemsystem oder Lautinventar beruht, kann als angewandte Linguistik betrachtet werden und ist als das Werk von hervorragenden Einzelpersonen erkennbar, z.B. Wulfila's Orthographie des Gotischen, des Notker Teutonicus und anderer Orthographien der althochdeutschen Dialekte.

Orthographie, d.h. das Schreibungssystem, die Verschriftung mit Normansatz, ist die Hauptquelle für die historische Phonologie. Das haben wir

oben (3. bis 6.) am Beispiel der deutschen Sprachgeschichte zu zeigen versucht. In der althochdeutschen Periode (750-1050) waren es vor allem Zeichenwahl für das Lautsystem und Zeichenwandel für den Lautwandel, welche die Forschung beschäftigt haben. Für die mittelhochdeutsche Periode (1050-1350) gibt uns die Germanistik ein gutes Beispiel für normalisierende Umschriftung der Schreibungen in den Handschriften (5.2 oben). Die Schriftdialekte des Frühneuhochdeutschen (1350-1720) zeigen auch charakteristische orthographische Variation, die durch Einwirkung der Fachleute sehr zurückgegangen ist. Die Nachkriegsgegenwart in deutschsprachigen Ländern mit ihrer Mundartwelle und allgemeinen Tendenz zu Entnormung und Fremdmustern sieht linguistische Pragmatiker am Werke (7. oben), die, vor allem unter dem Einfluß des Englischen, das deutsche Schriftbild ändern wollen.

BIBLIOGRAPHIE

- Antonsen, Elmer H. 1979. "The graphemic system of the Germanic Fuþark." In: I. Rauch and G.E. Carr (Hgg.) *Linguistic Method*, S. 287-297.
- Braune, Wilhelm u.a. 1969¹⁵. *Althochdeutsches Lesebuch*. Tübingen: Niemeyer.
- Chomsky, Noam und Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.
- Downing, J.A. 1967. *Evaluating the Initial Teaching Alphabet*. London: Cassell.
- Drosdowski, Günther. 1975. "Möglichkeiten und Grenzen einer Reform der Fremdwortorthographie," *Jahrbuch für Internationale Germanistik* 6, S. 8-19.
- Fishman, J.A. (Hg.) 1977. *Advances in the creation and revision of writing systems*. Den Haag: Mouton.
- Gelb, I.J. 1974. *A study of writing*. Chicago: University of Chicago Press.
- Hanna, Paul R. 1966. *Phoneme-grapheme correspondences as cues to spelling improvement*. Washington, D.C.: Office of Education.
- Hotzenköcherle, Rudolf. 1955. "Großschreibung oder Kleinschreibung? Bausteine zu einem selbständigen Urteil," *Der Deutschunterricht* 7, S. 30-49.
- Jellinek, Max Hermann. 1913f. *Geschichte der neuhighdeutschen Gramma-*

- tik von den Anfängen bis auf Adelung.* 2 Bd. Heidelberg: Winter.
- Keller, Rudolf E. 1978. *The German language.* New Jersey: Humanities Press.
- Kuhn, Wilfried. 1981. *Drei Prinzipien der Orthographieformierung. Ein Beitrag zum Problem der Verschriftung von Sprachen.* (KLAGE No. 7. Kölner Linguistische Arbeiten Germanistik.)
- Moulton, W.G. 1979. "Notker's 'Anlautgesetz'. In: I. Rauch und G.E. Carr (Hg.) *Linguistic Method*, S. 241-251.
- Müller, Johannes. 1882. *Quellenschriften und Geschichte des deutschsprachlichen Unterrichts bis zur Mitte des 16. Jahrhunderts.* Gotha. Neudruck: Hildesheim: Olms, 1969.
- Nerius, Dieter. 1975. *Untersuchungen zu einer Reform der deutschen Orthographie.* Berlin: Akademie-Verlag.
- Penzl, Herbert. 1971. *Lautsystem und Lautwandel in den althochdeutschen Dialekten.* München: Max Hueber Verlag.
- , 1972. *Methoden der germanischen Linguistik.* Tübingen: Niemeyer.
- , 1975. *Vom Urgermanischen zum Neuhochdeutschen. Eine historische Phonologie.* (Grundlagen der Germanistik 16). Berlin: Erich Schmidt Verlag.
- , 1982. "Zur Methodik der historischen Phonologie: Schreibung-Lautung und die Erforschung des Althochdeutschen," *Beiträge zur Geschichte der deutschen Sprache und Literatur* 104 (Tübingen), S. 169-189.
- , 1984. *Frühneuhochdeutsch.* (Langs Germanistische Lehrbuchssammlung, Band 9). Bern: Peter Lang.
- , 1984a. "Zum historischen Beweismaterial für Satzphonetik im Deutschen." In: *Fifth International Phonology Meeting. Discussion Papers*, S. 199-201 (*Wiener Linguistische Gazette. Supplemente. Beiheft 3.*)
- , 1985. "Sprachgermanisten und die Rechtschreibreform," *Internationales Jahrbuch für Germanistik*, 16:2, S. 74ff.
- Rauch, Irmengard und Gerald E. Carr. (Hgg.) 1979. *Linguistic Method. Essays in Honor of Herbert Penzl.* (Janua Linguarum. Series Maior, 79). Den Haag: Mouton Publishers.
- Voorwinden, N.Th.J. 1973. *Merigarto. Eine philologisch-historische Monographie.* Leiden, Universitaire Pers.
- Zehetner, Ludwig G. 1977. *Bairisch. Kontrastive Hefte für den Deutschunterricht.* Düsseldorf: Schwann.