Journal of English Linguistics

Editor
Charles F. Meyer
Applied Linguistics Program, University of Massachusetts–Boston
100 Morrissey Blvd., Boston, MA 02125-3393
meyer@cs.umb.edu

Associate Editors
Cynthia Bernstein
English Department, University of Memphis
Memphis, TN 38152-6176
cbernstn@memphis.edu
Anne Curzan
Department of English, University of Washington
Seattle, WA 98195-4330
acurzan@u.washington.edu
Pepi Leistyna
Applied Linguistics Program
University of Massachusets–Boston
100 Morrissey Blvd.
Boston, MA 02125-3393
pepi.leistyna@umb.edu

Review Editor
Claiborne Rice
English Department, University of Montevallo
Station 6420
Montevallo, AL 35115
ricecl@montevallo.edu

Editorial Assistant
Julian Jefferies

Senior Consulting Editors
Jan Aarts, University of Nijmegen (the Netherlands)
John Algeo, University of Georgia
W. Nelson Francis, Brown University
William A. Kretzschmar, Jr., University of Georgia
Donaldo Macedo, University of Massachusets–Boston
Virginia G. McDavid, Chicago State University
Fred C. Robinson, Yale University
Robert Stockwell, UCLA

Consulting Editors
Michael Adams, Albright College
Peter S. Baker, University of Virginia
J. K. Chambers, University of Toronto (Canada)
Richard Cureton, University of Michigan
Joan Hall, University of Wisconsin
Gregory Iverson, University of Wisconsin, Milwaukee
Roderick A. Jacobs, University of Hawaii
Merja Kytö, Uppsala University (Sweden)
Daniel Long, Tokyo Metropolitan University (Japan)
Christian Mair, University of Freiburg (Germany)
Donka Minkova, UCLA
Hans F. Nielsen, Odense University (Denmark)
Lee Pederson, Emory University
John Rickford, Stanford University
Edgar W. Schneider, Universität Regensburg (Germany)

For Sage Publications: Mark Gage, Gillian Dickens, Scott F. Locklear, Paul Doebler, and Jennifer Trone
CONTENTS

Contributors 293

Articles

Principles of Pattern Selection: A Corpus-Based Case Study
JOYBRATO MUKHERJEE 295

Is out of Always a Preposition?
BERT CAPPELLE 315

The Diphthongization of /ay/: Abandoning a Southern Norm in Southern Maryland
DAVID BOWIE 329

Yorkshire English Two Hundred Years Ago
MARÍA F. GARCÍA-BERMEJO GINER and MICHAEL MONTGOMERY 346

Reviews

Spreading the Word: Language and Dialect in America
By John McWhorter
PATRICIA CUKOR-AVILA 363

Qualitative Methods in Sociolinguistics
By Barbara Johnstone
ELLEN JOHNSON 367

Pathways of Change: Grammaticalization in English
Edited by Olga Fischer, Anette Rosenbach, and Dieter Stein
LAUREL J. BRINTON 372

The Continental Backgrounds of English and Its Insular Development until 1154
By Hans Frede Nielsen
FRED C. ROBINSON 376
Contributors

David Bowie is an Assistant Professor of Linguistics at Brigham Young University. He is the Director of the Utah Dialect Project.

Laurel J. Brinton is a Professor of English Language at the University of British Columbia. Her research interests include pragmatic markers, grammaticalization, and tense/aspect in the history of English.

Bert Cappelle is a Ph.D. student in Language and Linguistics (Germanic Languages) at the Catholic University of Leuven. He is doing corpus research within an interuniversity project (KULAK, KULeuven, Lille III), which is to result in a multivolume, descriptive grammar of contemporary English. His personal research interests focus on productive verb-particle patterns.

Patricia Cukor-Avila is an Associate Professor of Linguistics in the Department of English at the University of North Texas in Denton, Texas. Her work centers on the study of grammatical variation and change in rural Southern dialects, specifically African American English. Her longitudinal study of a rural Texas community has provided much of the data for presentations and articles concerning approaches to sociolinguistic fieldwork as well as documenting innovations in African American English. She is coeditor of The Emergence of Black English: Text and Commentary (1991) and author of several articles and book chapters. She is currently coauthoring a book with Guy Bailey entitled The Development of AAVE since 1850: The Evolution of a Grammar to be published by Cambridge University Press.

María F. García-Bermejo Giner is an Associate Professor of English Historical Linguistics at the University of Salamanca, Spain. Her main research interests are English dialectology in the nineteenth century and literary dialects since the Early Modern English period. She has also worked on nineteenth-century linguistic historiography and on colonial American English.

Ellen Johnson is an Assistant Professor of Linguistics at Berry College in Rome, Georgia. She wrote Lexical Change and Variation in the Southeastern United States, 1930-1990 and is currently researching language attitudes and cultural change.

Michael Montgomery is a Professor Emeritus of English at the University of South Carolina, where he taught from 1981 to 1999. He has recently completed a comprehensive dictionary of southern Appalachian English and is engaged in a long-term study of trans-Atlantic linguistic connections, especially to determine the influence of Ulster emigrants on American English.
Joybrato Mukherjee is an Assistant Professor of Modern English Linguistics at the English Department of the University of Bonn. His research interests include corpus linguistics, stylistics, text-linguistics, syntax, intonation, and EFL teaching. He is currently working on a corpus-based analysis of ditransitive verbs and their complementation patterns in present-day English.

Fred C. Robinson, FBA, is the Douglas Tracy Smith Professor Emeritus of English at Yale University. His specialties include English philology (of all periods) and Old English language and literature. His latest book is an edition (with Bruce Mitchell) of Beowulf (Oxford, 1998).

Paul E. Szarmach is the Director of the Medieval Institute at Western Michigan University and Professor of English and Medieval Studies. Formerly editor and now publisher of the Old English Newsletter, Szarmach has edited collections of essays on Old English prose and has written articles on Old English prose with special reference to Latin backgrounds. He has directed or codirected four NEH summer institutes or seminars.
Principles of Pattern Selection
A Corpus-Based Case Study

JOYBRATO MUKHERJEE
University of Bonn

Analyses of linguistic corpora have revealed that natural language is to a very large extent based on (semi-)preconstructed phrases. Drawing on corpus-based approaches to the description of such lexico-grammatical patterns in language use, the present study puts into perspective the question of why one and the same lexical item occurs in different patterns. The question of pattern selection (i.e., the analysis of factors that lead the language user to prefer a specific pattern in a given context) deserves further consideration. The present corpus-based case study is intended to illuminate this aspect of authentic language behavior.

Introduction: Pattern Grammar and Pattern Selection

Focusing on language patterns, Sinclair (1991, 110) formulates the idiom principle, which states

that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments.

Sinclair (1991, 110-21) himself gives several examples of lexical items that collocate with a restricted set of other lexical items (e.g., hard as in hard work, hard luck). Renouf and Sinclair (1991, 128-30) introduce the notion of collocational frameworks (e.g., “a/an + ? + of,” “many + ? + of”) as “discontinuous sequence[s] of two words” that “are highly selective of their collocates” in midposition. The selected collocates can be arranged in specific semantic groupings. For example, Renouf and Sinclair (cf. 1991, 136-37) show that the framework “an + ? + of” tends to co-occur with words from a limited number of semantic fields—for example, measurement (e.g., “an average of”) and relationship (e.g., “an enemy of”). That

AUTHOR’S NOTE: I would like to thank Jürgen Esser, Rolf Kreyer, and two anonymous referees for useful comments and suggestions. Furthermore, I wish to express my gratitude to Charles Meyer for criticism and encouragement. Both were needed.
language is in fact produced in and around patterns as lexico-grammatical units of meaning is not considered by traditional descriptive grammars, which are based solely on what Sinclair (1991, 109) calls the open-choice principle: “At each point where a unit is completed (a word, phrase, or clause), a large range of choice opens up and the only restraint is grammaticalness.”

An alternative to the traditional description of grammar as open choices is provided by the corpus-based analysis of more or less idiomatic, lexico-grammatical patterns. From a theoretical point of view, such a pattern-based grammar turns out to corroborate Sinclair’s (1991, 137) view that “when we have thoroughly pursued the patterns of co-occurrence of linguistic choice there will be little or no need for a separate residual grammar or lexicon.” Accordingly, Francis, Hunston, and Manning (1996, 1998) offer a corpus-based overview of all the patterns of all nouns, verbs, and adjectives that are attested in the Collins COBUILD English Dictionary. These pattern lists are supplemented with complementary lists indicating which lexical items share one specific pattern. In a wider context, Hunston and Francis (2000) introduce the concept of a comprehensive pattern grammar of the English language. The patterns of a word are defined by Hunston and Francis (2000, 37) as

all the words and structures which are regularly associated with the word and which contribute to its meaning. A pattern can be identified if a combination of words occurs relatively frequently, if it is dependent on a particular word choice, and if there is a clear meaning associated with it.

Patterns in this sense are thus based on collocational phenomena as well as on colligational co-selections (i.e., lexical and grammatical co-occurrences) (as defined by Sinclair 1991, 71 and Sinclair 1996, 85, respectively). From a complementary perspective, such a lexico-grammatical pattern is co-selected not only by specific words but by a more or less restricted range of words. Hunston and Francis (2000, 43) also state that “as a word can have several different patterns, so a pattern can be seen to be associated with a variety of different words. This is the opposite side of the coin.” It is obvious that in focusing on what is frequent in language use, the pattern grammar approach exemplifies a corpus-based and genuinely empirical grammar (i.e., an observation-based grammar) (cf. Aarts 1991, 44). Accordingly, Hunston and Francis draw on the large Bank of English corpus (comprising 329 million words in June 1999).

In a similar vein to the Collins COBUILD English Dictionary, Hunston and Francis (2000, 131) use a comparatively small inventory of iconic symbols to represent patterns: for example, “v-link ADJ for n to-inf” symbolizes a pattern that is frequently found with the adjective difficult as in “it’s quite difficult for us to do...”
Table 1 shows how Hunston and Francis (cf. 2000, 135) visualize the meaning of the pattern itself.

As one can see, the different positions in the pattern are attached to different semantic aspects of the process of evaluation, which is evoked by the use of difficult in this pattern. To show that a pattern itself also co-selects a specific range of words, the words of a pattern are subsumed into so-called meaning groups. The pattern “Vn n,” for example, chooses verbs from five meaning groups if the verb is ditransitive: (1) “giving someone something, or refusing to do so”; (2) “doing something for someone”; (3) “communicating something to someone”; (4) “giving someone a benefit or a disadvantage”; and (5) “verbs concerned with feeling and attitudes.” On the other hand, if the verb in the pattern “Vnn” is complex-transitive, there is only one meaning group—namely, “putting something into a category” (Hunston and Francis 2000, 87-90).

This article is clearly to be seen in the tradition of the pattern grammar approach to the English language. However, it is not the description of the patterns of a word as such that lies at the heart of this study. Rather, this study addresses a somewhat more basic question: if a lexical item occurs in different patterns, which principles and factors make the language user choose a specific pattern? The lemma PROVIDE is very suitable for a case study of this problem of pattern selection since it has four formally different patterns that are semantically similar in that they select equivalent thematic roles:

1. V n n (“provide someone something”)
2. V n with n (“provide someone with something”)
3. V n for n (“provide something for someone”)
4. V n to n (“provide something to someone”)

All patterns refer to the use of PROVIDE for an action in which a provided entity is passed to an affected entity (in the sense of intended or actual recipient) (cf. Quirk et al. 1985, 696-697). It turns out that Hunston and Francis’s (2000) representation of patterns has to be modified for two reasons. First, the affected entity may also be inanimate, as is the case with forum in the following example:

Table 1 shows how Hunston and Francis (cf. 2000, 135) visualize the meaning of the pattern itself.

As one can see, the different positions in the pattern are attached to different semantic aspects of the process of evaluation, which is evoked by the use of difficult in this pattern. To show that a pattern itself also co-selects a specific range of words, the words of a pattern are subsumed into so-called meaning groups. The pattern “V n n,” for example, chooses verbs from five meaning groups if the verb is ditransitive: (1) “giving someone something, or refusing to do so”; (2) “doing something for someone”; (3) “communicating something to someone”; (4) “giving someone a benefit or a disadvantage”; and (5) “verbs concerned with feeling and attitudes.” On the other hand, if the verb in the pattern “V n n” is complex-transitive, there is only one meaning group—namely, “putting something into a category” (Hunston and Francis 2000, 87-90).

This article is clearly to be seen in the tradition of the pattern grammar approach to the English language. However, it is not the description of the patterns of a word as such that lies at the heart of this study. Rather, this study addresses a somewhat more basic question: if a lexical item occurs in different patterns, which principles and factors make the language user choose a specific pattern? The lemma PROVIDE is very suitable for a case study of this problem of pattern selection since it has four formally different patterns that are semantically similar in that they select equivalent thematic roles:

1. V n n (“provide someone something”)
2. V n with n (“provide someone with something”)
3. V n for n (“provide something for someone”)
4. V n to n (“provide something to someone”)

All patterns refer to the use of PROVIDE for an action in which a provided entity is passed to an affected entity (in the sense of intended or actual recipient) (cf. Quirk et al. 1985, 696-697). It turns out that Hunston and Francis’s (2000) representation of patterns has to be modified for two reasons. First, the affected entity may also be inanimate, as is the case with forum in the following example:
Consequently sport provides the forum for the development of leadership skills and relationships, . . . (FLOB F05 62)

Thus, it is useful to rephrase the meaning of the patterns to be discussed here so that animate as well as inanimate affected entities are covered. Second, the two noun groups are indexed in the following patterns so that the affected entity (n₁) and the provided entity (n₂) in each pattern are clearly distinguished:

(1′) V n₁ n₂ ("provide someone/something something")
(2′) V n₁ with n₂ ("provide someone/something with something")
(3′) V n₂ for n₁ ("provide something for someone/something")
(4′) V n₂ to n₁ ("provide something to someone/something")

It should be noted in passing that some instances of PROVIDE formally resemble one of these patterns but are in fact quite different in meaning and have, therefore, not been taken into account. This pertains, for example, to the preposition for introducing a time adjunct or the preposition to introducing not the affected entity but rather the subject matter (cf. Quirk et al. 1985, 709-10):

(6) It provides an excellent guide to inter-agency co-operation . . . (FLOB H09 110)

In the following section, the frequencies of the four patterns in several corpora will be presented and discussed. It will be shown in particular that the frequencies are relatively stable across corpora. On this basis, some largely monocausal hypotheses as to the selection of patterns will be rejected. Pattern selection turns out to be much more complex a phenomenon based on many factors rather than one explanatory principle only.

**Distribution of the Patterns in Five Corpora**

The first step of the analysis involved searching for all occurrences of PROVIDE in the patterns (1′) to (4′) in four standard 1 million-word corpora: the Lancaster-Oslo/Bergen Corpus (LOB) of written British English with texts from 1961, the Freiburg LOB Corpus (FLOB) with texts from 1991-1992, the Brown Corpus (BROWN) of written American English with texts from 1961, and the Freiburg Brown Corpus (FROWN) with texts from 1991-1992. These parallel corpora contain roughly the same genres and were compiled according to the same standards. Their analysis thus allows hypotheses regarding linguistic change and variation in present-day English to be tested (cf. Mair 1997, 195-98): whether there is regional variation between British and American English (in the 1960s and in the 1990s) and diachronic change in the two varieties (including mutual influences). In addition,
the much larger British National Corpus (BNC) was analyzed (and the results will be discussed later in this section).

Table 2 shows how many occurrences of PROVIDE are associated with each of the four patterns (1') to (4'), as well as the corresponding percentages in relation to the total number of occurrences of PROVIDE in each 1 million-word standard corpus.7

The figures in Table 2 reveal neither significant diachronic change nor significant regional variation. The distribution of the patterns is very stable across the four standard corpora.8 This observation is at odds with the hypothesis put forward by Hunston and Francis (2000, 97):

Although *provide* is typically used with the pattern *V n with n* (‘provide someone with something’), there are a handful of occurrences in the Bank of English of ‘provide something to someone’ (the pattern *V n to n*), presumably by analogy with *give*.

Thus, they suggest a diachronic change in the use of patterns (e.g., in the case of PROVIDE) due to a general tendency to use the same pattern with words that have a similar meaning (e.g., GIVE). Hunston and Francis (2000, 97) admit, however, that their hypothesis cannot be verified as there are no adequate corpora available for the English language with which the large Bank of English corpus could be compared: “For this reason, it is not possible to say for certain that a particular, infrequent usage is new just because it occurs in a later (and larger) corpus but not in an earlier (and smaller) one.” In the light of the pattern distribution across the four standard corpora, though, the hypothesis of an ongoing diachronic change appears to be questionable, to say the least. Furthermore, one could argue that it is not the pattern “*V n with n*” but rather the pattern “*V n, for n*,” with which PROVIDE is typically used.

It has often been suggested that the pattern “*V n, n*” of PROVIDE is restricted to American English (cf. Quirk et al. 1985, 1210). Accordingly, this pattern is only attested in the BROWN and FROWN corpora. Since it only occurs marginally, however, it is doubtful whether it is truly relevant to a discussion of pattern selection concerning PROVIDE. It does, quite obviously, not even provide a genuine alternative to American language users. One could hypothesize that the use of the pattern “*V n, n*” is not so much a matter of regional variation but of idiosyncratic preference.

Largely capitalizing on Hawkins’s (1994) research, Rohdenburg (1996, 149) introduces a *complexity principle* “to account for the distribution of competing constructions involving different degrees of explicitness.” Concerning the four semantically similar patterns of PROVIDE, the complexity principle would lead to a preference of the less explicit pattern “*V n, n*” to the other, grammatically more ex-
### TABLE 2
Relative Frequencies of Four Patterns of PROVIDE in Four Standard Corpora

<table>
<thead>
<tr>
<th>Corpus</th>
<th>Word-Form</th>
<th>V n_1 n_2</th>
<th>V n_1 with n_2</th>
<th>V n_1 for n_2</th>
<th>V n_1 to n_2</th>
<th>Σ Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOB (BrE, 1961)</td>
<td>provide</td>
<td>0</td>
<td>4</td>
<td>34</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>provides</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>provided</td>
<td>0</td>
<td>0.0%</td>
<td>11</td>
<td>6.0%</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>providing</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>FLOB (BrE, 1991-1992)</td>
<td>provide</td>
<td>0</td>
<td>16</td>
<td>44</td>
<td>17</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>provides</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>32</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>provided</td>
<td>0</td>
<td>0.0%</td>
<td>3</td>
<td>5.9%</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>providing</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>BROWN (AmE, 1961)</td>
<td>provide</td>
<td>3</td>
<td>19</td>
<td>51</td>
<td>13</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>provides</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>35</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>provided</td>
<td>0</td>
<td>0.6%</td>
<td>6</td>
<td>6.9%</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>providing</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>3</td>
</tr>
<tr>
<td>FROWN (AmE, 1991-1992)</td>
<td>provide</td>
<td>0</td>
<td>14</td>
<td>46</td>
<td>26</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>provides</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td>34</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>provided</td>
<td>2</td>
<td>0.7%</td>
<td>9</td>
<td>5.9%</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>providing</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>12</td>
<td>8</td>
</tr>
</tbody>
</table>

**NOTE:** LOB = Lancaster-Oslo Bergen Corpus; FLOB = Freiburg LOB Corpus; BROWN = Brown Corpus; FROWN = Freiburg Brown Corpus.
licit patterns in cognitively less complex environments. However, the marginal frequency of the less explicit pattern (i.e., the pattern without a preposition) makes it impossible to apply this principle to the pattern selection under discussion.

Furthermore, in all corpora, all word-forms are attested in all the patterns (albeit to different extents). The only exception is the pattern “V n₁ n₂.” This, however, does not come as too much of a surprise, considering its very few occurrences. As far as the other, substantially more frequent patterns are concerned, there are no categorical “sense restrictions on specified word-forms” (Esser 2000, 97) that would result in exclusive associations between one word-form of PROVIDE and one specific pattern and its meaning.

Apart from the fact that regional variation and diachronic change cannot account for the versatility of patterns under scrutiny in this study, it is reasonable to assume that it is, in general, not one factor alone that needs to be considered. Rather, I would suggest that the selection of a specific pattern is based on different and conflicting principles. The language encoder prioritizes one or several of these principles and thus prefers a certain pattern to others in a given context. In my view, the corpus data in Table 2 might point to the fact that the set of those principles has not changed across time and does not differ significantly between American and British English. This is not to say that the corpus data are not relevant to the discussion of the factors responsible for the selection of one pattern or another. Rather, a functional approach to the principles at work should be able to account for the general tendencies in all the standard corpora: the pattern “V n₂ for n₁” is much more common than the patterns “V n₁ with n₂” and “V n₂ to n₁.”

Before turning to a systematic analysis of the principles responsible for pattern selection, it should be noted that the distribution of the patterns in the four standard corpora can also be found in larger corpora. I also searched the 100 million-word BNC for the four patterns given in (1’ to 4’). It attests 55,021 occurrences of the lemma PROVIDE in total: 22,312 instances of provide (40.6 percent), 8,360 instances of provides (15.2 percent), 17,003 instances of provided (30.9 percent), and 7,346 instances of providing (13.3 percent). Due to the limitations of the software tool SARA (cf. Aston and Burnard 1998, 55), the solution had to be thinned, and I did this proportionally concerning the relative quantities of the word-forms. Thus, I let SARA compile random selections of all the word-forms with 406 downloads of provide, 152 downloads of provides, 309 downloads of provided, and 133 downloads of providing, resulting in a random selection of 1,000 concordance lines. Table 3 presents the findings.

Although differing in absolute frequencies, the relative quantities are not statistically different from the ones in the four standard corpora. That is to say, the standard corpora, albeit relatively small and restricted to written language, are reliable for the purpose at hand in light of the BNC data, which are taken from a larger corpus with written and spoken texts.
<table>
<thead>
<tr>
<th>Corpus</th>
<th>Word-Form</th>
<th>V n₁ n₂</th>
<th>V n₁ with n₂</th>
<th>V n₂ for n₁</th>
<th>V n₁ to n₁</th>
<th>Σ Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNC (BrE, 1960-1993)</td>
<td>provide</td>
<td>0</td>
<td>35</td>
<td>75</td>
<td>24</td>
<td>PROVIDE</td>
</tr>
<tr>
<td></td>
<td>provides</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td></td>
<td>provided</td>
<td>0</td>
<td>0.0%</td>
<td>9</td>
<td>6.1%</td>
<td>148</td>
</tr>
<tr>
<td></td>
<td>providing</td>
<td>0</td>
<td>0.0%</td>
<td>9</td>
<td>6.1%</td>
<td>25</td>
</tr>
</tbody>
</table>

NOTE: BNC = British National Corpus.
Principles of Pattern Selection: A Functional Analysis

In this section, I confine myself to those patterns that occur in significant frequencies in all corpora: “V n₁ for n₂,” “V n₁ with n₂,” and “V n₁ to n₂.” For obvious reasons, the pattern “V n₁ n₂” is left out of consideration since it only occurs sporadically in the two American English corpora.

So far, the focus has been on the semantic similarity of the patterns under discussion. They are considered as logically equivalent in that they select the same thematic roles and refer to the same type of action (i.e., passing a provided entity to an affected entity). Although all kinds of affected entities have been searched for in the corpus analyses, corpus data suggest that there is a clear restriction on the pattern “V n₁ with n₂.” The affected entity in this pattern is usually animate (cf. Quirk et al. 1985, 314) as in the following examples:

(7) ... your study will provide you with the knowledge that is generally accepted ... (BNC EEB 179)
(8) ... providing teachers with scientific resources and project materials ... (FLOB H12 216)
(9) ... the placenta is unable to provide the baby with oxygen ... (FLOB F31 164)
(10) ... which is to provide the nation with food of the highest quality ... (LOB H10 65)
(11) ... is quite willing to provide the police with a doughnut, a cup of coffee, or a meal ... (FROWN G62 188)

Examples (10) and (11) illustrate that animate gender classes also include items that may be either personal or nonpersonal: nation and police are collective nouns and can potentially be used to refer to “the individuality within the group” (Quirk et al. 1985, 316). Also, animal nouns and names of countries are often used with a personified reference. Considering the broad dichotomy between animate items (in such a wider sense) and inanimate items, the affected entity in the pattern “V n₁ with n₂” is usually animate, whereas no such clear-cut preference can be found in the pattern “V n₁ for n₂.” Table 4 shows that this difference holds true for all five corpora.

Obviously, the pattern “V n₁ with n₂” tends to co-select a much more restricted lexis in the n₁-position than the pattern “V n₂ for n₁” does. In other words, the former pattern has a more restricted meaning than the latter. This also explains why the pattern “V n₂ for n₁” prevails in all corpora: it turns out to be semantically more flexible as far as the affected entity is concerned.

In these two patterns, the affected entities (and provided entities as well) are bound to different positions. It is therefore reasonable to have a look at some pragmatic principles that are responsible for word order variation. In particular, I would like to refer to the principle of end-focus and the principle of end-weight. Quirk
et al. (1985, 1357) state that “it is common to process the information in a message so as to achieve linear presentation from low to high information value,” a phenomenon covered by the principle of end-focus. Thus, the decision on which pattern to choose may be guided by this principle: if the affected entity is to be focused, the pattern “V n<sub>2</sub> for n<sub>1</sub>” will be preferred; if the provided entity has a higher information value, the pattern “V n<sub>1</sub> with n<sub>2</sub>” will be selected. In the following examples, the pattern “V n<sub>2</sub> for n<sub>1</sub>” is chosen so that the affected entity can be placed in end-focus position. In example (12), it is obvious that what a cow provides—namely, “milk”—is rather thematic, whereas the fact that “everyone in the locality” can be provided with milk is the new information. In example (13), “for Wiltshire” is in end-focus position for reasons of syntactic parallelism. In examples (14) to (17) as well, the range of persons affected by the action turns out to be most important (consider the specification of the affected persons by means of premodifying adjectives, postmodifying relative clauses, or participle constructions).

(12) A white cow used to provide milk for everyone in the locality, . . . (BNC BMT 420)
(13) . . . then he’ll put more money into Wiltshire and provide more police officers for Wiltshire. (BNC JS9 87)
(14) Should the government directly provide education for the children who want public education? (BROWN J48 1950)

<table>
<thead>
<tr>
<th></th>
<th>V n&lt;sub&gt;1&lt;/sub&gt; with n&lt;sub&gt;2&lt;/sub&gt;</th>
<th>V n&lt;sub&gt;2&lt;/sub&gt; for n&lt;sub&gt;1&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Animate Affected Entity</td>
<td>Inanimate Affected Entity</td>
</tr>
<tr>
<td>LOB (all)</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>91.6</td>
<td>8.4</td>
</tr>
<tr>
<td>FLOB (all)</td>
<td>27</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>84.4</td>
<td>15.6</td>
</tr>
<tr>
<td>BROWN (all)</td>
<td>32</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>91.4</td>
<td>8.6</td>
</tr>
<tr>
<td>FROWN (all)</td>
<td>31</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>91.1</td>
<td>8.8</td>
</tr>
<tr>
<td>BNC (random)</td>
<td>57</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>93.4</td>
<td>6.6</td>
</tr>
</tbody>
</table>

NOTE: LOB = Lancaster-Oslo Bergen Corpus; FLOB = Freiburg LOB Corpus; BROWN = Brown Corpus; FROWN = Freiburg Brown Corpus; BNC = British National Corpus.
(15) The museum has received a special award for providing outstanding facilities and services for disabled people. (FLOB E03 215)

(16) To solve the elder-care problem, he would provide “choices” for old people who still have a lot of money. (FROWN A14 158-159)

(17) Ricky wanted a license mainly because it provided an identification card and a degree of cover for someone seeking work. (FROWN G69 103-104)

If the provided entity is to be focused, the pattern “V n₁ with n₂” tends to be chosen. In examples (18) to (20), the affected entity is a personal pronoun and clearly represents given information. In (21), the noun “world” is inferable from world-knowledge; in (22) and (23), “these nations” and “the drivers” respectively have already been mentioned before. Thus, in all cases, it is the provided entity that represents the rhematic item and is accordingly placed in final focus position by using the pattern “V n₁ with n₂.”

(18) He could even provide me with a gun if I needed one. (BNC G15 3229)

(19) It also conveniently provided me with straight edged divisions of the remaining space. (BNC CN4 417)

(20) ... an arty individual, whose specialty is the American boy and who adopts a 10-year-old to provide him with fresh idea material. (BROWN C04 0110-0120)

(21) ... the EC must be strengthened to provide the world with a counter-weight to the USA. (FLOB F17 168)

(22) ... conserve biodiversity in developing nations any more than it can in the developed nations. What research can do, however, is provide the people and the leaders of these nations with information that may help them to improve their lives, ... . (FROWN H06 195-198)

(23) ... to compensate drivers for any apparent risks in trucking. In addition, it is quite possible that firms provided the drivers with greater safety resources . . . (FROWN J41 104-106)

While the principle of end-focus refers to the arrangement of items according to their information values, the principle of end-weight operates at the level of complexity of constituents (cf. Quirk et al. 1985, 1362): there is a general tendency to place heavier (i.e., more complex) constituents after lighter constituents. The articulation of two separate principles does, however, not rule out that they often come into operation along with each other “[s]ince the new information often needs to be stated more fully than the given (that is, with a longer, ‘heavier’ structure)” (Quirk et al. 1985, 1361). In some of the previously mentioned examples, for example, (14) and (19), the principle of end-focus obviously coincides with the principle of end-weight. Focusing on the relevance of weight to the pattern selection, the pattern “V n₂ for n₁” tends to be used if the affected entity is more complex than the pro-
vided entity. The pattern is used because it allows the affected entity to be placed in final position, as in examples (24) and (25): “1,700 mentally abnormal people who are judged to . . .” and “the game of musical chairs they played” are considerably heavier than “care” and “the music,” respectively.

(24) The special Hospitals Broadmoor, Rampton and Ashworth Hospitals provide care for 1,700 mentally abnormal people who are judged to . . . (BNC FYW 1096)

(25) The lute also provided the music for the game of musical chairs they played, . . . (BNC G3M 963)

In examples (26) and (27), on the other hand, the provided entity (“a spiritual core around which they can orbit” and “bases from which nuclear weapons can be used,” respectively) is more complex than the affected entity. Thus, the pattern “V n₁ with n₂” is chosen since it places the more complex entity toward the end.

(26) . . . the New Age provides seekers with a spiritual core around which they can orbit, . . . (FLOB D15 20)

(27) . . . and provide the Americans with bases from which nuclear weapons can be used. (LOB B23 19)

To summarize, it has been suggested that the choice of one of the two patterns “V n₂ for n₁,” and “V n₁ with n₂,” is influenced by (1) the animacy of the affected entity and (2) the pragmatic principles of end-focus and end-weight. Unlike the pattern “V n₁ with n₂,” the pattern “V n₂ for n₁,” is much more frequent as it is not restricted to animate affected entities. If a free choice of patterns is possible, pragmatic principles may influence the pattern selection.

The third pattern of interest here is “V n₁ to n₁.” Note that it resembles the pattern “V n₂ for n₁,” concerning the order of elements. Examples (28) and (29) reveal that the two patterns even seem to be interchangeable in very similar contexts:

(28) . . . shall provide technical assistance and funds to States for training for public safety officials . . . (FROWN H15 22-23)

(29) In carrying out the requirements to provide technical assistance and funds for training, . . . (FROWN H15 151-152)

Despite the same order of elements in the two patterns, the quantitative corpus analysis shows that the “V n₂ to n₁” pattern is less common than the “V n₂ for n₁” pattern. Thus, it is reasonable to suggest that there are also general restrictions on the pattern “V n₂ to n₁.” It is interesting that this pattern turns out to co-select a restricted set of provided entities. That is to say, it is used in combination with a more or less seman-
tically restricted lexis in the n2-position. The attested items in \textit{LOB}, \textit{FLOB}, \textit{BROWN}, \textit{FROWN}, and \textit{BNC} include, for example, the following:

(30) aid, assistance, answer, boost, care, challenge, contribution, grant, help, impetus, incentive, information, input, protection, sanctuary, service, solace, solution, stimulus, subsidy, support, treatment, value

This list highlights the conclusion drawn from earlier corpus analyses that PROVIDE has a fairly positive semantic prosody. The notion of semantic prosody refers to the “consistent aura of meaning with which a form is imbued by its collocates” (Louw 1993, 157), which in case of PROVIDE is positive concerning the provided entities (cf. Stubbs 1995, 26). Perhaps this overall rating can be diversified by taking into account the different patterns of PROVIDE. Whereas the pattern “V n1 to n2” almost exclusively selects “pleasant” entities in the n2-position, the pattern “V n2 for n1” also co-selects a number of provided entities that may be regarded as “neutral” (e.g., background, basis, context, framework, and structure). Considering, however, the comparatively small random selection taken from the \textit{BNC}, this differentiation between more positive and less positive semantic prosodies of PROVIDE, depending on the selected pattern, is only a tentative suggestion. To verify this hypothesis, a more exhaustive corpus analysis would no doubt be required.

Many of the lexical items used in the n2-position in the pattern “V n1 to n2” quite generally co-select the preposition \textit{to} even (and especially) when they are not used in this pattern. In other words, many items in the n2-position have a pattern themselves that could be represented as “N to n.” This conclusion can be drawn from the pattern information in the \textit{Collins COBUILD English Dictionary}. Almost all n2-items in the pattern “V n1 to n2,” which I found in the five corpora, are habitually associated with the pattern “N to n” according to the definitions, examples, or patterns given in the \textit{Collins COBUILD English Dictionary}. These items include the ones listed under (30). The only lexical item that occurs frequently and does not have the pattern “N to n” is \textit{advice}. Disregarding this exception, there seems to be a general and strong tendency to choose the pattern “V n1 to n2” whenever a lexical item that generally co-selects the preposition \textit{to} is to be used in the n2-position of the pattern “V n2 for n1” (being the default case in terms of relative quantity). The following three examples illustrate this principle:

(31) ... thus providing a more effective \textit{challenge to} independent services. \textit{(FLOB G76 196)}

(32) ... it provides the only realistic \textit{solution to} the problems of race relations ... \textit{(LOB D17 84)}

(33) ... governments are able to provide local \textit{subsidy to} local firms or individuals ... \textit{(FROWN H05 153)}
Challenge, solution, and subsidy are examples of nouns that are usually associated with the pattern “V to n” according to the Collins COBUILD English Dictionary. Obviously, the overall pattern of PROVIDE is adapted to the usual pattern of the lexical item in the n2-position by changing the preposition, leading to a congruence between the pattern and the n2-item.

It should be noted, though, that there are quite a few instances of lexical items that are used in both patterns of PROVIDE (i.e., “V n2 for n1” and “V n2 to n1”—for example, assistance, funds, help, impetus, incentive, and service). This makes it clear that patterns describe frequent regularities in authentic language use but no rigid rules in terms of, say, grammaticality. The flexibility and creative potential of language notwithstanding, there are well-defined principles that account for the majority of cases found in corpora. I have tried to outline the relevance of some of these factors to the selection of the patterns under discussion as far as PROVIDE is concerned. In this, three important aspects have been tackled: (1) some patterns may select a restricted lexis in either the n1- or the n2-position; (2) the attempt to give priority to one or the other pragmatic principle might lead the language user to prefer a specific pattern; and (3) some lexical items are habitually associated with a specific pattern of their own resulting in choosing a pattern that reconciles this small-scale pattern with the large-scale pattern of, say, the governing verb.

Table 5 brings the different aspects of pattern selection together and gives a systematic account of the factors that lead the language user to prefer a specific pattern of PROVIDE to others. For lack of sufficient corpus data on the pattern “V n1 n2,” only the remaining three patterns are taken into account.

It is quite clear now why the pattern “V n2 for n1” is more frequent than the other two patterns in all corpora. It is neither restricted in terms of the gender class of n1 nor in terms of the preposition selected by n2. This pattern is therefore very flexible
and can be used with virtually all affected and provided entities: it is the default case of pattern selection.

It goes without saying that there may be other principles and/or factors at work too. For example, one could argue that the pattern “V n₁ to n₁” puts into operation the “Animated First Principle,” which states that “the NP which is most animated will precede NPs which are less animated” (Tomlin 1986, 102). This principle, however, originates in typological research comparing the basic word orders of different languages. All the same, its language-specific implication for PROVIDE is captured by the notion of animacy restriction (cf. Table 5). Furthermore, the principle of end-focus is clearly linked to the Prague School given-before-new view of the typical English sentence. That is to say, the focus is said to be at the end in the unmarked or neutral order of elements. On the other hand, Givón (1983) puts forward the contradictory principle of task urgency. Mair (1990, 35) points out that neither approach has been empirically verified so far and that, more important, the “postulation of ‘basic’ orders—be it ‘new-before-given’ or ‘given-before-new’—is bound to result in gross over-simplification.” In my view, it is therefore better to keep to the principle of end-focus for the time being and regard it as a dynamic principle that the speaker may or may not follow for specific communicative reasons. Whether it really refers to the unmarked (i.e., most frequent) information structure in English remains to be seen. Finally, Hawkins (1994) suggests a cognitive principle according to which language users tend to choose an order of elements that unveils the immediate constituent structure as soon as possible to facilitate the language decoder’s processing task. The principle of end-weight correlates with this principle: if light constituents precede heavy constituents, the immediate constituent structure of the clause is opened up earlier than it would be in the reverse order. Taking these considerations into account, I assume that many other general principles of word-order variation are implicitly covered by the factors summarized in Table 5. Moreover, those three factors alone are able to explain, in general, the quantitative trends of pattern distribution in the corpora under discussion.

Concluding Remarks: Envisaging a Systematic Approach to Pattern Selection

Quite obviously, the outline of principles of pattern selection offered in the present study is suggestive rather than exhaustive. Nevertheless, the corpus-based case study of PROVIDE and its occurrence in four formally different, but semantically similar, patterns hopefully broadens the perspective opened up by the pattern grammar approach. Corpus linguistics should not (and will certainly not) confine itself to a mere description of all the patterns a given word frequently co-selects. Since most lexical items choose a variety of patterns and, more important, several logi-
cally equivalent patterns, future research in this field should further investigate the principles that underlie the selection of patterns in authentic contexts. This will surely deepen our understanding of how language users make effective and efficient use of the patterns that they have at their disposal. It will then be possible to explain in functional terms why a word occurs in various patterns in the first place. Furthermore, the identification of such principles would certainly enhance the impact that corpus linguistics has already had on language teaching. Not only could language learners be provided with information about which patterns to use with which words, but they could also learn which specific pattern of a word to use in a given context. On the basis of 100 randomly selected concordance lines from the Bank of English, Hunston and Francis (2000, 130-32), for example, identify 21 patterns of DIFFICULT, some of which are very similar regarding their thematic role assignment (e.g., “find it ADJ to-inf” and “it v-link ADJ for n to-inf”) (cf. Table 1). It should be interesting to carry out further corpus analyses to describe lexical restrictions and other factors that may determine which pattern is more likely to be selected under specific contextual circumstances. Finally, this kind of information about pattern selection could usefully be integrated into future approaches to usage-based models of language competence as, for example, suggested by Langacker (1987, 494), who envisages a “nonreductive approach to linguistic structure that employs fully articulated schematic networks and emphasises the importance of low-level schemas” and that focuses on “the actual use of the linguistic system and a speaker’s knowledge of this use.” The principle-guided use of patterns on the basis of functional criteria for pattern selection is certainly a good candidate for inclusion in such a usage-based model.

On the basis of the corpus data on PROVIDE, as presented and discussed in this article, at least three aspects need to be considered: (1) the meaning of the pattern itself in that it may co-select a more restricted set of lexical items than other patterns; (2) the influence of pragmatic principles, if different patterns have different arrangements of thematic roles and can thus be selected according to, say, the principles of end-focus or end-weight; and (3) the congruence between a large-scale pattern and the selected lexical items that themselves may habitually be associated with small-scale patterns of their own. It is beyond reasonable doubt that with regard to other lexical items and their patterns, further principles need to be suggested. However, the fact that language users have a fundamental freedom of choice in their pattern selection despite all principle-guided and empirically accessible probabilities cannot be exaggerated. Even a comprehensive account of all principles of pattern selection will always cover regularities of actual language use only and fail to state rigid rules. Nonetheless, it is no doubt this very description of what is probable rather than what is possible (cf. Kennedy 1998, 270) that makes corpus linguistics so fascinating and relevant to the understanding of real language.
Notes

1. It should be noted that it is not only in English linguistics that such an integrated lexico-grammar has been envisaged. In French linguistics, Gross (e.g., 1993, 1995) suggests a *local grammar* referring to the patterns of a word in given contexts. This local grammar provides an important framework for corpus processors capable of automatically identifying the patterns of a word in large texts. INTEX is such a corpus processor that is based on a local grammar of French (cf. Silberztein 1993 and Faron 1999, who give a detailed account of INTEX).

2. However, Francis (1993, 139) insists that there is a fundamental difference between the corpus-driven approach that she favors and the corpus-based grammar suggested by Aarts in that the latter explicitly allows for intuition-based methods in an observation-based grammar. Because of the notorious unreliability of intuition, Francis is correct in postulating that “data comes first.” However, there will always be a need for reference to the linguist’s intuition even when working with corpora, particularly in the functional interpretation of the raw material provided by corpus data.

3. As the adjective is under discussion, its symbolic code is capitalized. If a specific lexical item is part of the pattern (e.g., the preposition *for*), it is given in italics. The symbol “v-link” denotes the link verb, “n” a noun group, and “to-inf” a clause introduced by a to-infinitive.

4. As Stubbs (1995, 24) summarizes, a “*lemma* is a dictionary head-word, which is realised by various *word-forms.*” Throughout this article, lemmas are cited in upper case and word-forms are italicized.

5. It is quite surprising that this pattern, which turns out to be by far the most frequent one of the four patterns under discussion, is not indicated in the *Collins COBUILD English Dictionary* (Sinclair 1995, 1324-25), which only gives the following patterns: “V n,” “V n with n,” “V that,” and “V P n” (the latter referring to the construction “provide for someone/something”). On the other hand, it is not surprising that the pattern “V n n” is not included in this dictionary since it occurs only marginally in American English (see Table 2).

6. This terminology neglects the fact that corpus linguistic research has revealed a clearly negative semantic prosody of *affect:* “Things are usually badly or adversely affected” (Stubbs 1995, 45). *Provide,* on the other hand, has a positive prosody (cf. Stubbs 1995, 26).

7. As there is a substantial divergence between the corpora concerning the total figures of PROVIDE, measuring the relative quantities of the patterns appears to be reasonable.

8. This is also corroborated by statistical procedures such as the chi-square ($\chi^2$) test, which allows for an estimation of whether frequencies in corpora differ
from each other to a significant extent (cf. Oakes 1998, 24-29). From the application of the chi-square test to the figures in Table 2 (British vs. American corpora, corpora from the 1960s vs. corpora from the 1990s, Lancaster-Oslo Bergen Corpus [LOB] vs. Freiburg LOB Corpus [FLOB], Brown Corpus [BROWN] vs. Freiburg Brown Corpus [FROWN]), the general conclusion can be drawn that the observed frequencies do not differ significantly from the expected frequencies at a significance level of 0.05. For example, the application of the chi-square test to LOB and FLOB gives a $\chi^2$ of 3.19 (with 3 degrees of freedom), which is clearly below the significance level of 7.82. With regard to BROWN and FROWN, $\chi^2$ is 3.84.

9. Sinclair (1991, 8) states that “each distinct [word-]form is potentially a unique lexical unit.” Drawing on this idea, Esser (2000, 97) defines a “lexical linguistic sign” as “the union of a single sense and a set of medium-independent, abstract grammatical word-forms. The set may include all or only a subset of the possible morphological forms.” This concept could be applied to associations between specific word-forms and the patterns in which they are used. However, with regard to the patterns of PROVIDE to be discussed in this article, there are no such restrictions on any of the word-forms.

10. Consider also examples (28) and (29).

11. As a matter of fact, this reservation also holds for the previously suggested relevance of pragmatic principles that may, of course, be infringed or violated for specific reasons.

References


Is out of Always a Preposition?

BERT CAPPELLE
University of Leuven/Kortrijk, Belgium

The collocation out of is usually called a preposition. It obviously differs from simplex prepositions like against, at, like, from, with, and so on in being made up of two separate words. It is often labeled, therefore, a “complex preposition,” a “compound preposition,” or a “two-word preposition.” This is also what sets it off from prepositions like into, onto, and upon, which are written as single words although they historically derive from two words as well (in + to, on + to, up + on).

Hardly anyone would question the analysis of out of as a preposition in combinations like She acted out of pity or It was something out of the ordinary. In other contexts, however, there might be some doubt as to whether out of should really be treated as a single category. The alternative would then be to analyze out of as a combination of an adverbial particle (out) and a preposition (of). A case in point is the expression to run out of time in the following sentences, which are taken from the fully tagged and parsed ICE-GB (1998) corpus:

(1) On present policies, it will run out of time. (ICE-GB W2E-010-106)
(2) You must be running out of time. (ICE-GB S1A-052-043)

Since (1) and (2) involve the same expression, one would expect to find a similar syntactic analysis attached to them in ICE-GB. This is not the case. The syntactic analyses in ICE-GB come in the form of syntactic trees. Figures 1 and 2 reproduce the (slightly simplified) trees that are given for (1) and (2).

In Figure 1, out of is treated as a complex preposition. In Figure 2, out and of are analyzed as belonging to two separate phrases, an adverbial phrase (AVP) headed by the adverbial particle out and a sister prepositional phrase with of. In fact, the expression to run out of time has three occurrences in ICE-GB. Apart from the sentences already mentioned, there is also the following sentence:

(3) No no, I’m not I’m not running out of time [unclear words] (ICE-GB S1A-052-054)

The expression to run out of time has received yet a third parsing in this utterance (see Figure 3).
In Figure 3, the *of*-prepositional phrase is parsed as a postmodifier of *out*, which is the head of a complex adverbial phrase.² If we look at how other occurrences of *out of* in ICE-GB have been parsed, we notice that the labeling assigned to (1) is by far the commonest, with 467 instances. The analyses as for (2) and (3) are considerably less frequent, with 7 and 3 instances, respectively. These low frequencies might simply be due to some minor inconsistency in the postediting of the output of the parsing program, which generated more than one phrase structure tree for most sentences. Still, what remains worth our interest is the fact that multiple parse trees were at all present in the output. The observation that the posteditors did not consistently reject some of these multiple parsings further highlights the unclear status of *out of*.

The analyses for (2) and (3) crucially differ from the analysis for (1) in their labeling of *out* as a separate word rather than as a part of a complex preposition. The central problem could therefore be phrased as follows: *is out in out of always prepo-
sitional, or can it arguably be treated as an adverbial particle? In other words, does *out* always construe with *of* to form a prepositional phrase, or does it sometimes belong to a verb-particle construction?

Note that this problem does not exclusively arise with idiomatic combinations like *to run out of money*. The following sentences are again extracted from the *ICE-GB* corpus. Both contain the sequence *out of the water*, which describes a path in physical space. In (4), *out of* is treated as a two-word preposition, while in (5), *out* is labeled a separate adverb.

(4) Uhmm as soon as I took my leg out of the water it fell straight open. (*ICE-GB S1B-066-026*)

(5) . . . she didn’t realise anything was wrong with her leg until the blood started obviously after she’d come out of the water. (*ICE-GB S1B-067-151*)

In the next section, I will list the arguments that can be adduced in favor of *out of* as always being analyzed as a preposition and show how each of the arguments fails to be entirely convincing. I will conclude that there are no persuasive reasons to reject analyzing *out of* as a particle + preposition.

### Reasons to Treat *out of* as a Preposition

*Out* and *of* Are Syntactically Adjacent

Perhaps the most compelling argument to treat *out of* as a complex preposition is the fact that it cannot be split up. If the first part (*out*) were a particle, it should be
able to occur in front of a direct object of a transitive verb–particle combination, but this is not borne out by the facts. Compare the grammaticality of (6b) with the ungrammaticality of (7b):

(6) a. Desmond took the pistol out.
   b. Desmond took out the pistol.
(7) a. Desmond took the pistol out of his pocket.
   b. * Desmond took out the pistol of his pocket.

Since *out cannot move in (7), one might conclude that it cannot be a particle. However, it has been noted as a general tendency that if a transitive verb–particle combination is followed by a directional prepositional phrase, the particle is placed just before the prepositional phrase (cf. Gries 1999, 110; his example):

(8) a. He put the junk down onto the floor.
   b. He put down the junk onto the floor.

This tendency can be observed with all sorts of directional prepositional phrases, not just those that express a true direction (i.e., a destination) but also those that express a source or a trajectory:

(9) a. . . . he played a ball through to Lineker. (CBD)
   b. . . . he played through a ball to Lineker.
(10) a. Ella’s voice boomed so loudly that Autumn moved the receiver away from her ear slightly. (CBD)
    b. ? Ella’s voice boomed so loudly that Autumn moved away the receiver from her ear slightly.
(11) a. The freedom to spend, to retain wealth and to pass that inheritance on down through the generations . . . (CBD)
    b. ? . . . and to pass on that inheritance down through the generations . . .
    c. ? . . . and to pass on down that inheritance through the generations . . .

Clearly, the addition of a directional prepositional phrase (be it a destination, a source, or a trajectory) seriously hinders the positioning of a particle to the left of the object. It might well be the case that if this prepositional phrase is introduced by *of, the tendency to place the particle to the right of the object is even stronger than with other directional prepositional phrases. (The reason for this could be that of cannot express direction in itself but needs “support” from the particle *out.) Whether this strong clustering of *out and *of is sufficient reason to see *out then as an integral part of this prepositional phrase is still questionable.
Out of Can Be Pied-Piped

Just like real prepositions (e.g., into) and unlike clear combinations of particles + prepositions (e.g., out from), the sequence out of can occur in the so-called pied piping pattern (i.e., with a relative or interrogative pronoun immediately following it):

(12) ... deepening the hole out of which the economy has to climb? (ICE-GB W2C-008-082)
(13) ... the traps into which the new property-based local tax could easily fall. (ICE-GB W2E-009-058)
(14)* She nodded toward Patrick, who had found a toy out from which a foam ball popped.

Moreover, out cannot stay behind on its own after the verb, again like inseparable prepositions and unlike clear combinations of particles + prepositions:

(15)* ... deepening the hole of which the economy has to climb out?
(16)* ... the traps to which the new property-based local tax could easily fall in.
(17) She nodded toward Patrick, who had found a toy from which a foam ball popped out. (CBD)

This is clear evidence that out of is a preposition and not a concatenation of a particle plus a preposition, so it seems.

However, it must be noted that the sequence out of which is significantly less frequent than into which, both in ICE-GB (1 vs. 10 occurrences) and in CBD (31 vs. 233 occurrences). It could be objected, though, that into is simply much more frequent than out of, not just in combination with which but in any position. This is true: in ICE-GB, into occurs 1,528 times, whereas out of only occurs 477 times, as mentioned above. In CBD too, there are more occurrences of into (81,207) than of out of (30,467). But even so, the proportion of out of which is significantly smaller than the proportion of into which (p < .05 when a chi-square test was applied to the CBD figures). This observation somewhat tones down the conclusion that out of behaves completely analogously to the preposition into.

Furthermore, the ungrammaticality of (15) does not necessarily prove that out is not a particle—it might simply prove that the preposition of cannot always be pied-piped. Indeed, of is semantically more empty than, for example, from, which might account for its different syntactic possibilities.

Into Is Not in_{ps} + to_{prep}, So out of Is Not out_{ps} + of_{prep}

The preposition into is not simply the sum of the particle in plus the preposition to. As Declerck (1976, 10) puts it,
The meaning of into the room is not identical with that of in plus that of to the room. If this were the case, I went into my father would be synonymous with *I went into my father, which is incongruous. Similarly, onto is not equivalent to on plus to.4

If into cannot be equated with in to, there seems to be no reason to assume that the sequence out of should be split up into out + of either, with out being a particle and of a preposition.

However, this reasoning is not very sound: if into is not the same as in to, the reverse must be true as well (i.e., the sequence in to is not an alternative spelling of the preposition into). Consider the following example. (Here and throughout the text, brackets are used to show syntactic structure, not to insert added material in the quotes.)

(18) He came across as good and simple and true, which helped the checks [roll in] [to Project Ararat]. (Julian Barnes, A history of the world in 10 1/2 chapters)

The combination in to in (18) cannot be pied-piped, so it is not a preposition. By contrast, since to is an uncontroversial preposition, it follows naturally that it can be pied-piped. Compare (19a) and (19b):

(19) a. * In to which project did the checks roll? 
    b. To which project did the checks roll in?

Also, the sequence on to is not always (if sometimes) a spelling variant of the preposition onto:

(20) Let’s [move on] [to the next thing]. (ICE-GB S1B-071-251)

This means that the above argument can be reversed: if the string in to is not a preposition and if the string on to is not always a preposition, there is no reason to assume that out of is always a preposition either. Stated positively, if in to is clearly a concatenation of a particle + preposition and if on to is often one as well, then it is possible to analyze out of as a particle + preposition too.

Out of Is a Fixed Combination

Out of is such a common collocation that it is best treated, one might say, as a single (complex) word, just like on top of or in addition to (see Kjellmer 1994 for frequency information on collocations of this nature).
However, while of and to are the only possible prepositions after on top or in addition, respectively, of is not the only preposition that can occur after out. Declerck (1976, 11) gives the following combinations, which do not exhaust all possibilities:

(21) a. He threw the ball out of the kitchen.
    b. He threw the ball out to the waiting players.
    c. He threw the ball out across the lawn.
    d. He threw the ball out into the garden.
    e. He threw the ball out onto the roof.
    f. He threw the ball out over the wall.
    g. He threw the ball out under the trees.

Counting all these combinations (out of, out to, out across, out into, etc.) as complex prepositions would lead to an explosion of prepositions. It is therefore better to analyze them as “free” combinations of particles and prepositions. Of course, out of is a much more common combination than, say, out across, but this does not in itself warrant a different analysis.

**Out of Is a Unique Niche Filler**

Out of is claimed to be the only grammatical expression available for the prepositional meaning “negative destination” from within an area or a containing volume (Quirk et al. 1985, 674). More specifically, the addition of of to out has often been observed to be indispensable:

(22) He walked out of (not out) the house. (Bolinger 1971, 34)
(23) He walked out of the room. (*out the room) (Lindner 1984, 253, n. 3)
(24) The children gathered some bluebells before they went out of/*out the wood and returned home. (Quirk et al. 1985, 678)
(25) *John threw the cat out the house. (O’Dowd 1998, 130)

Since it is the case, first, that the gap in He walked . . . the house is filled by prepositions (e.g., towards, into, around, through, etc.), and since it appears to be the case, second, that the opposite meaning of into can be expressed by out of and only out of, it must be the case that out of is a preposition.

However, out (without of) actually does occur as a preposition to express motion/removal from within a container-like space:

(26) . . . as you [go] [out the house that they’re working on], just stick your hand out again. (ICE-GB S1A-028-223)
(27) . . . he [goes] [out the room] for two minutes . . . (ICE-GB S1A-047-173)
(28) . . . that she [’d climbed] [out the pool] . . . (ICE-GB 1SB-076-151)
(29) And they [get] [out the car] and one of the thieves goes to the Queen, “Can we have all your money?” (ICE-GB S1A-041-311)

(30) ... the elderly priest sitting on a bench, [teasing some of her drying vaginal mucus] [out his groin fur], with a wistful, almost spiritual expression on his greying muzzle. (Will Self, Great Apes)

This usage of *out* is admittedly rather infrequent, but nonetheless, its occurrence implies that *out of* is not the sole element to fulfill this particular prepositional task.

If *out* can function as a preposition, we might wonder whether we are still compelled to call *out of* also a preposition. After all, the gap in *He walked . . . the house* is not only filled by prepositions but also by combinations of adverbial particles and prepositions (e.g., *away toward, off into, back through*, etc.). Should we not reserve a tagging for *out of* that is different from the tagging of *out*?

Since, as we have just shown, *out* is a real preposition and since, of course, *of* is a preposition, one such tagging could be *out*_{prep} + *of*_{prep}. This tagging might seem no more than an alternative notation for [*out of*]_{prep} but is in fact similar to the tagging of combinations like *from under*. However, there is an important difference between *John threw the cat out of the house* and *John pulled the cat from under the chair*. In the first sentence, *of the house* can be omitted so that *out* is the last word, but in the second sentence, *from* cannot be the last word. Therefore, the only serious alternative for [*out of*]_{prep} is *out*_{pri} + *of*_{prep}.

*Out of* Conjoins with *into*

The fact that *out of* can be conjoined with the (uncontroversial) preposition *into* pleads for the view in which *out of* is a true preposition too:

(31) . . . an account of how Britain bumbled into and out of the war . . . (CBD)

(32) The transponder was the signaling device required by law on all aircraft flying into and out of the US. (CBD)

(33) You will need to climb into and out of Zodiacs for excursions ashore. (CBD)

However, the collocation *in and out of* is much more frequent than *into and out of*:

(34) She went in and out of hospitals where she was given electroconvulsive therapy and heavy doses of medication. (CBD)

(35) Her breath hissed in and out of her nostrils, the only sound she could hear. (CBD)

(36) Dubious characters sidled in and out of liquor stores and discreetly gaudy prostitutes lurked in the shadows. (CBD)
In these examples, *in* cannot possibly be a preposition (e.g., *... went in hospitals*). Also, it does not associate with *of* (e.g., *... went in of hospitals*). Its use, therefore, must be adverbial. The ease with which *in* and *out* combine suggests that *out* is felt to have the same categorial status as *in*. By consequence, we should dissociate *out* from *of* and assign adverbial status to it as well.

**Conclusions and Further Reflections**

When the sequence *out* + *of* + NP indicates a literal or metaphorical path after verbs like *go, come, walk, run, take, grab*, and so on, there are three possible syntactic analyses:

(37) a. [\_x [ prep out of ] NP]
    b. [\_y [ Adv out ] [ prep [ prep of ] NP]]
    c. [\_z [ Adv out ] [ prep [ prep of ] NP]]

In (37a), *out of* is analyzed as a complex preposition introducing a prepositional phrase. The alternative is to consider *out* as a separate word—namely, an adverbial particle, as in (37b) and (37c). The first analysis is the one most commonly used, and this practice is reflected in the way *out of* is tagged in the *ICE-GB* corpus.

In this respect, it is interesting to see how the grammatical labels in the *Cobuild* dictionary differ in the two editions. In the first edition (1987), the user could find in the entry *out* such diverse labels as the following:

(38) a. ADV AFTER VB
    Someone else I know has just come out of hospital. . . . I had been out of university a year. . . . Once out of high school she started singing in night clubs.
    b. ADV AFTER VB, OR PREP
    She opened a lacquered box and took out a cigarette. . . . She had the key out and was fumbling at the door. . . . He got out a book and read. . . . She tore several sheets of paper out of the back of the book. . . . Could you take it out of the fridge for me? . . . The fossils are carefully chipped out of the rock.
    c. ADV AFTER VB/N: IF + PREP THEN USU of
    She rushed out of the house. . . . The lift doors opened and they stepped out into the empty foyer. . . . He was already on his way out. . . . I got up to see him out. . . . She’s just got out of bed. . . . It dropped out of the sky.
    d. ADV AFTER VB, OR ADV + of
    . . . a week-end resort about twenty-five miles out of town. . . . I don’t want to live any further out. . . . I can’t wait to get out of Birmingham.
What is confusing here is not only the diversity of the grammatical labels used but also the apparent mismatches with the examples that ought to correspond to them and the inconsistency of their application. For example, why does the label in (38a) not specify “+ of” if all three illustrative sentences apparently contain out + of? Why is out considered a preposition in some of the examples in (38b)—presumably those examples in which of follows—but an adverb possibly followed by of in (38c) and (38d)? In the second, revised edition (Cobuild 1995), the editors have tried to bring order to this chaos by making a strict distinction between “adverb uses” and “preposition uses.” Under the first heading are listed only uses of out without of. Under the second heading are listed all and only the uses of out followed by of.

In this article, though, I have tried to show that there is nonetheless something to be said for the view in which out can be tagged as an adverb (i.e., an adverbial particle) even when it is followed by of. I have done so by examining several arguments that could be given in favor of the analysis as in (37a)—[[PP [Prep out of] NP]]—and showing that none of them is fully persuasive. The conclusion is not that those who call out of a complex preposition in, for example, the following sentences are wrong:

(39) a. I usually cook out of tins and jars and packets. (ICE-GB S1A-059-087)
   b. He tumbled out of bed . . . (ICE-GB W2F-001-022)

Rather, the conclusion that I want to put forward is that those who wish to call out a particle in at least (39b) cannot convincingly be shown to be wrong either.

I would like to end this study with two more reflections. First of all, the question whether out in sentences like (39b) is part of a prepositional phrase (out of bed) or part of a verb-particle combination (tumble out) also arises with some other complex prepositions, such as away from, ahead of, up to, off of, and so on. Mutatis mutandis, it could be defended that these collocations should not always unequivocally be called complex prepositions either. Let us consider the following pair of sentences:

(40) a. Slice the top off each tomato and empty carefully. (CBD)
   b. You want me to slice little pieces off of you? Is that what you want? (CBD)

In (40a), off is clearly prepositional, but in (40b), this is less clear.

Second, the problem discussed in this study only arises within syntactic frameworks that strictly adhere to phrase structure. In phrase structure grammars, lexical items have one and only one tagging in a given sentence, and nodes cannot have more than one immediate mother node. The following “tree” diagram (see Figure 4) is anomalous in phrase structure grammar, even if it might intuitively capture the structure of the sentences in hand.
(The PP could also have been a daughter node of the VP instead of a sister node, but this is beside the question.) Crucially, in this analysis, the word *out* is tagged as both an adverb and a preposition. It therefore resembles a so-called *prepositional adverb* (cf. Quirk et al. 1985, 713-16), the difference being that the prepositional complement is not ellipted but overtly expressed. The result is a lexical item belonging to two phrases, which is not problematic in phrase structure grammar as long as one phrase is totally *included* within the other. Here, however, the lexical item in question belongs to partially *overlapping* phrases:

(41) she [VP'd come {PP out} of the water]

While this constitutes a problem in phrase structure grammar, there might not be a problem at all in some other syntactic frameworks. Dependency Grammar, for example, does not allow nonterminal nodes, and consequently there are no phrases to be classified separately from their heads (Hudson 1980).
So, on a final and perhaps comforting note, the problem—as usual—lies in the eye of the beholder.

Notes

1. *ICE-GB* (1998) is the British component of the *International Corpus of English (ICE)* containing one million words of contemporary spoken and written British English collected and annotated at the Survey of English Usage, University College London. For more information on *ICE-GB*, see http://www.ucl.ac.uk/english-usage/ice-gb/index.htm.

2. It might seem strange that the sequence *out of time* (whatever the exact analysis of its internal structure) is not integrated within the VP in any of the trees, the more so since these words are linked with the verb *run* to form an idiomatic expression and do not denote a path in physical space. The *ICE-GB* corpus, however, does not stick to binary branching conventions (e.g., $S \rightarrow NP \ VP$) and distinguishes idiomatic combinations like (1) through (3) from “free” combinations like *she came out of the pool* by adding to the relevant words (i.e., *out of*) the feature *phrasal* versus *general*, respectively.

3. *CBD* stands for *CobuildDirect*, an online service for accessing a 56 million-word corpus of modern English language texts, both written and spoken. It contains mainly British English, but American and (to a smaller extent) Australian English are also represented. The texts in this corpus have been selected from the Bank of English, launched in 1991 as a joint project of COBUILD (a division of HarperCollins publishers) and the University of Birmingham.

4. With *onto*, this orthographic test does not work, since *onto* is often written as two words even if its use is clearly prepositional:

So he has bought a laptop computer, on to which each day’s transactions are downloaded . . . (*CBD*)

5. At least of is claimed to be indispensable when reference is made to the containing source itself. Lindner (1984, 253, n. 3), Quirk et al. (1985, 678), and O’Dowd (1998, 130) note that *out* can appear without *of* if there is reference to an *opening* in the containing source through which there is movement (e.g., out of the door, out of the window, etc.). Bolinger even claims that the addition of *of* is ungrammatical in that case:

He walked out (not out of) the door. (Bolinger 1971, 34)

While in American English, indeed, *out the door/window* is much more frequent than *out of the door/window* (162 vs. 27 occurrences in *CBD*), Bolinger’s judgment does not reflect actual language use for British English. Only in spoken British
English does this usage of *out* without *of* appear to be dominant (67 vs. 37 occurrences in CBD), but not in written language, where *out the door/window* is less common than *out of the door/window* (103 vs. 162 occurrences in CBD).

6. As is well-known, sound emission verbs like *screech* and *roar* can also be followed by directional *out* (*of NP*) (e.g., *The car screeched out of the parking lot*). In fact, many verbs that do not per se code a motional event can be thus construed (e.g., *He sat up in his tree and nearly sneezed himself out of it*).

7. The head of an arrow points to a word’s dependent. For instance, in the example, the noun *water* is represented as dependent on the determiner *the*, following common practice in Dependency Grammar (and in current mainstream generative grammar). The determiner *the* depends on *of*, which depends on *out*, and so on. It could be defended that the particle *out* is not dependent on the verb *come*, but rather the other way round. Particles have often been claimed to be “more central than the verb” (Declerck 1977, 311). Discussing combinations like *look/glance/run/flick/skim/leaf/rifle/thumb through*, Bacchielli (1993, 58) argues that “the particle *through* ends up expressing the basic idea of the action and so acquires a verbal force, whilst the preceding verbs . . . are confined to a modal, or instrumental function and thus become premodifications of *through*.”

**References**


The Diphthongization of /ay/
Abandoning a Southern Norm in Southern Maryland

DAVID BOWIE
Brigham Young University

Although /ay/ in Ireland reduced to [ə] and [Ø] was reported as far north as New York state relatively early in the twentieth century (Linguistic Atlas of the Middle and South Atlantic States 2000), monophthongal /ay/ is generally thought of as a marker of Southern American English by linguists and nonlinguists alike. The difference between Northern and Southern levels of monophthongal /ay/ is, in fact, great enough that Thomas (1997) was able to use a decrease in the rate of monophthongal /ay/ among urban Texan Anglos as evidence for massive dialect contact with non-Texans.

Bowie (2000, 52) notes in passing a decrease in the rate of monophthongal /ay/ over the past century of apparent time among natives of Waldorf, a medium-sized community in Southern Maryland, but does not propose a reason for this pattern. A closer look at the data on /ay/-monophthongization in Southern Maryland allows us to test whether the roots of this development in exurban Maryland can be traced to the same cause as similar developments in other speech communities.

Waldorf

Waldorf is a town of 51,324 (1995 estimate) located twenty-three miles south-southeast of Washington, D.C., at the northern edge of Charles County, Maryland. Whether Waldorfians have a “Southern” or a “Northern” accent is actually a topic of conversation in the town, and several informants mentioned /ay/-monophthongization as an example of Southern speech (generally by offering a demonstration such as “It’s [t.aːm] to eat”). Previous dialectological studies give some insight into why the Southernness or Northernness of Waldorf English is such a salient issue for Waldorfians—Waldorf is located at or near a dialect border. Dialectologists disagree on the exact position of the line separating the South and South Midland dialect regions in Maryland, as Figure 1 shows. This map shows the approximate position of the South-South Midland dialect isogloss according to

AUTHOR’S NOTE: I would like to thank Hikyoung Lee, Don Norton, and two anonymous reviewers for comments on earlier versions of this work.

© 2001 Sage Publications

329
Kurath and McDavid (1961) and Kretzschmar et al. (1993); Waldorf is marked by the oval at the northern edge of Charles County. Note that according to Kurath and McDavid (1961), Waldorf lies clearly within the Upper South dialect region (specifically, in the Virginia Piedmont region), while according to Kretzschmar et al. (1993), Waldorf lies in an undefined border area between the Southern and South Midland regions.¹

Because Waldorf is an unincorporated municipality, precise historical data on the population of the town are difficult to come by, but the population of Charles County as a whole has increased since 1950 at a faster rate than national and state averages, as Figure 2 shows (U.S. Bureau of the Census 1995, 2001a, 2001b, 2001c). Also, the population growth rate of northern Charles County (where Waldorf is located) appears to have been increasing at a rate exceeding that of Charles County as a whole since the middle of the twentieth century (Edelen et al. 1976; Potyraj 1994), which means that the rate of population increase in Waldorf has been extremely high. It should be noted that what local residents generally think of as Waldorf proper has never held very many people at all. Historically, the town was solely a commercial center surrounded by farms, and now the situation remains the same, but with fewer farms and the addition of housing developments throughout the area. Therefore, it should be kept in mind that what is referred to as Waldorf in this article is actually more properly the Waldorf area.

This rapid growth in Waldorf and Charles County populations has come largely from increases in employment opportunities with the federal government, as well
as private employment (largely white-collar) in Washington, D.C. and its suburbs. As a result, the increase in the area’s population has come from all over the United States, leading to a great deal of dialect mixture. Such intense dialect mixture has been found to lead to dialect leveling in several urban locations (see, among others, Kerswill and Williams 1994, 11; Kerswill 1994, 70; 1996a, 241; 1996b, 298; Thomas 1997, 328), and the situation in Waldorf provides a chance to look at the situation in a semi-rural exurban location.

Whether Waldorf is historically Southern or South Midland, then, two things should be particularly noted to give context. First, Waldorf is at or near the northern edge of the Southern dialect region, and so Waldorfians have long had easy access to models of non-Southern linguistic norms. Second, this access to non-Southern models has become even easier and more commonplace since World War II.

**Method and Data**

Taped interviews were conducted between the fall of 1997 and the summer of 1999 with twenty-five white middle-class natives of Waldorf, fourteen female and eleven male. These interviews contained a mixture of elicitation of linguistic tokens (primarily through minimal pair elicitation and word lists), elicitation of demographic data, and a more casual section directed toward the production of personal narratives on the part of the subject. The interviews ranged from about thirty minutes to two hours, depending on the talkativeness of the subject. The elicitation of tokens of /ay/ was not a goal of the interviews, so most of the tokens of /ay/ the subjects produced came during periods of narrative or other casual speech.
The tapes of the interviews were then subjected to impressionistic coding, with the first ten minutes of each interview excluded from review. All the tokens of /ay/ that an individual produced after that first ten minutes (up to a maximum of 200 tokens) were coded, with certain exceptions. The exceptions were that instances of the words I and my were each capped at 10 percent of an individual’s total tokens, and all /ay/ followed by /y/ were excluded from analysis. Eleven speakers yielded the maximum of 200 tokens, and only three (Beatrice, Dean, and Warren) yielded less than 100, with Beatrice producing the fewest, 52; a grand total of 4,090 tokens were collected and coded. Tokens of /ay/ were coded as either “glide present” or “no glide”; “no glide” was defined as /ay/ pronounced without any discernable glide at all. Of the 4,090 tokens, 396 (9.68 percent) were pronounced as monophthongs.

Along with the social factors of age and sex (class and race being constant among the speakers), tokens were coded for stress (primary or secondary), style (casual interview conversation, narrative, elicited word, and word list), syntactic environment (subject noun, other noun, verb, adjective, adverb, interjection, and other), following morphological boundary, and following sound. Following Hazen’s (2000) suggestion, not only was the nature of the following sound noted (nasal, liquid, voiceless obstruent, /w/, vowel, pause, and filled pause) but also whether the following sound was tautosyllabic or not. Finally, instances of the word I (along with the contractions I’m, I’d, I’ll, etc.) and the word like when used as a discourse particle (but not as, for example, a verb) were coded as separate categories, in case those particular words acted strangely. The data were then subjected to both an apparent time analysis and a multivariate analysis using VARBRUL. To perform the multivariate analysis, the data were first checked to determine whether any factor groups showed interactions that would introduce errors into a VARBRUL analysis (Sankoff 1988, 992). This resulted in some factor groups being combined (as described later in this article) before the VARBRUL analysis itself was begun.

**Monophthongization of /ay/ in Apparent Time**

Significant differences were found when the overall results for /ay/-monophthongization rates were broken down only by social factors (sex and year of birth of the speakers). As Figure 3 shows, the rate at which individuals produce monophthongal /ay/ shows a clear decrease over apparent time. The decrease over time for men provides a very good fit to an exponential curve, with \( R^2 = 0.9593 \). The fit for the women is not quite as good, at \( R^2 = 0.7734 \), but this could be seen as simply reflecting the fact that one often finds more variation among women as a group than men as a group (Labov 2001). In any event, it appears clear that this is a change in progress, with women generally leading men in the change by about a generation’s worth of time (note that lower percentages on the graph in Figure 3 re-
flect further progress in the trend away from monophthongal /\textit{ay}/. This separation of the sexes is unsurprising, as it is the pattern that one would ordinarily expect from a situation in which a speech community is trending away from a linguistic feature that the community has begun to recognize and stigmatize (Labov 1990, 244).

**Interaction between Following Sound and Syllable Position**

The effect of following sound on the monophthongization of /\textit{ay}/ is shown in Figures 4 and 5. Figure 4 illustrates the effect of the following sound without regard to syllable position. Unsurprisingly, the more sonorant the following sound, the more likely it is that /\textit{ay}/-monophthongization occurs (with the exception of pauses, which do not have an easily assigned sonorance value, and vowels). Figure 5, however, presents this information broken down more finely, in that tautosyllabic following sounds are separated from nontautosyllabic ones. (Note that following vowel, /\textit{w}/, filled pause, and pause are all inherently nontautosyllabic.) This finer breakdown supports Hazen’s (2000) claim that syllable position and not just sonorance affects monophthongization of /\textit{ay}/. The exact relationship between syllable position and sonorance of the following sound is not entirely easy to describe, however, as tautosyllabic liquids and voiced obstruents favor /\textit{ay}/-monophthongization more than their nontautosyllabic counterparts, while the reverse is true for nasals and voiceless obstruents.

Because a VARBRUL analysis is designed to deal with interactions between factors, it might seem logical to use VARBRUL to explain the results shown in Fig-
ures 4 and 5. However, this is not possible because of the interaction between following sound and syllable position—they are not independent factors. The lack of independence among these factors cannot be explained away simply by an analysis that would have a syllable boundary weaken the effect of the following sound, bringing its effect closer to the overall mean (9.68 percent monophthongization).

This analysis could work for following liquids (where the rate of monophthongization is 24.83 percent for all following liquids but 14.46 percent when a syllable boundary follows) and voiceless obstruents (where a syllable boundary raises the rate from 1.31 percent overall to 4.07 percent). However, a syllable break increases the rate of monophthongization for following nasals slightly away from the mean.
(19.03 percent overall to 20.15 percent with a syllable break), and a syllable break reduces the rate for following voiced obstruents from slightly above the mean to even farther below the mean (10.01 percent overall to 5.88 percent). In addition, the environments of following vowel, /w/, filled pause, and pause necessitate following syllable breaks due to simple phonetic realities. This fact alone, even absent other observations, poses a problem for any VARBRUL analysis of the data that would keep following sound and following syllabicity as separate factors—the interaction between following sound and syllabicity is outside of the sort of interaction that VARBRUL analysis can explain, as they are not independent factors. That is, even though a VARBRUL analysis is designed to take interactional effects into account, it can be successful at this only if the effects are independent (Sankoff 1988, 992). As the effects of following sound and following syllabicity are not independent in this data set, a VARBRUL analysis cannot deal with them separately, and so the factors of following sound and following syllabicity were combined into a single factor group for the multivariate analysis, following Sankoff’s (1988, 993) suggestion.

**Multivariate Analysis**

A multivariate analysis of the factors influencing /ay/-monophthongization gives further insight into the specifics of the way the change is occurring in this speech community. For all speakers as a group, no factor was found to force the monophthongization of /ay/, and only two factors were found to consistently prevent it: /ay/ occurring in interjections and in the word like used as a particle. These cases were collapsed into broader “other” categories to allow the analysis to proceed.

A step-up step-down analysis of the data found one factor insignificant, which was then eliminated from further analysis: I (and contractions containing I) versus other words. Another step-up step-down analysis was then run, and all remaining factors were found to be significant. The stepping-up process added the factor group of following sound first, then age of speaker, then sex of speaker, syntactic function, syllable stress, and finally style; the inclusion of these factor groups was confirmed in the step-down portion of the analysis. Although not necessarily a perfect reflection of the relative strength of the effect of each of these factors, the order in which factor groups are added in a VARBRUL step-up analysis can be taken as a rough guideline of their relative significance from highest to lowest. Therefore, the VARBRUL analysis results for each of the factors are discussed below in that order.

The following sound was found to have an extremely strong influence on monophthongization of /ay/. A look at the returns for this factor group from the VARBRUL analysis (shown graphically in Figure 6) generally shows the sorts of results one would expect, given previous research on /ay/-monophthongization.
That is, following liquids very strongly favor monophthongization of /əɪ/ with nasals and voiced obstruents also favoring monophthongization. On the other hand, voiceless obstruents disfavor monophthongization, in some cases more strongly than liquids favor it (depending on syllabicity). Particularly worth noting is the distance between the effect of following tautosyllabic liquids (a VARBRUL weight of 0.938) and that of following tautosyllabic voiceless obstruents (0.030), reflecting the strength of this factor.

It is worth noting that the interaction between following sound and syllabicity, noted earlier, is affirmed by this result, though with some clarity added. The effect of liquids and voiced obstruents favoring monophthongization is lessened when they are accompanied by a syllable break, while the effect of nasals remains nearly unchanged by syllabicity. The effect of following voiceless obstruents disfavoring monophthongization is lessened considerably when there is an intervening syllable break, to the extent that a nontautosyllabic voiceless obstruent is nearly neutral in its effect on monophthongization. A summary generalization is that nasals act differently than these other sounds, in that a syllable break has a miniscule effect on nasals’ tendency to favor monophthongization. Why following nasals should be so different from other sounds is an issue that merits further investigation.

VARBRUL weights for the ages of speakers are shown in Figure 7. Although not quite as striking as the results for following sound, they still show a very large effect. The data in Figure 7 as a measure of changing tendencies away from /əɪ/-monophthongization over apparent time support the earlier claim, based on the data in Figure 3, that this is a change in progress. In particular, it should be noted that even though the second oldest age group (those born between 1920 and 1939)
rather strongly favor the monophthongization of /ay/, they do not favor it nearly as strongly as the oldest age group (those born before 1920). As will be discussed later, this fact has ramifications for any attempt to find the causes of the community’s abandonment of /ay/-monophthongization.

Men generally favor the monophthongization of /ay/, and women disfavor it, with VARBRUL factor weights of 0.618 for men and 0.410 for women. This supports the claim made earlier that women are leading this change since the direction of the change is away from monophthongal /ay/. Such a result is certainly not unexpected—the trend away from monophthongal /ay/ is a change in progress, and as noted earlier, Waldorfians are aware of this feature of Waldorf English, so one would expect women as a group to lead the change (Labov 1990, 244).

The syntactic function of the words /ay/ occurs in does not have as wide a range of VARBRUL weights as following sound and age of speaker, but it does have a noticeable effect; VARBRUL weights for this factor group are shown graphically in Figure 8. Nouns favor /ay/-monophthongization somewhat (nonsubject nouns favoring it slightly more strongly than subject nouns), as do adverbs and verbs; adjectives and all other words disfavor it a bit more strongly than other categories favor it. (The “other” category includes such cases as prepositions, determiners, and ambiguous cases.) Perhaps the most noteworthy issue to stress in relation to this is that syntactic function really does have an independent effect, which means that (for example) the word fine used as a noun would more likely be pronounced with a monophthong than the same word used as an adjective.

Although syllable stress does have a significant effect on /ay/-monophthongization in Waldorf, its effect is not large. Secondary stress favors /ay/-monophthongization more than primary stress (VARBRUL weights of 0.590 and 0.487, respectively), but this effect is perhaps simply the result of English phonetics. That
is, unstressed vowels are often reduced in English, and so the favoring of monophthongal /ay/ in syllables with secondary stress may well be the result of a reduction of the diphthong through removal of the glide.

The style of the utterance has very little effect, with one exception. The VARBRUL weights for style (as Figure 9 shows) are close to 0.500 for narrative, casual interview, and elicited word styles. Reading from a word list, however, strongly disfavors /ay/-monophthongization, with a VARBRUL weight of 0.257. Although one might expect reading from a word list to result in a move toward the standard diphthongal /ay/ both because of formality and possible reinforcement of the diphthongal form through spelling, it is somewhat surprising that no particular style pushed strongly away from the standard. One interesting application of this fact, though, is that for this variable in this speech community, any method of data collection other than reading tasks may well result in equally reliable data.

The Actuation Problem

Standing in the background in nearly all studies of language change is the actuation question: Why did this change ever start? In this case, though, the actuation question is especially intriguing because here we have what appears to be the reversal of a local norm in favor of a norm from somewhere else, the sort of change found by Thomas (1997, 310) in urban Texas and Boberg and Strassel (2000, 118) in Cincinnati, Ohio. (It is also worth noting that this change, like the one Thomas reports on, involves a change toward a “standard” model.) This is not an unheard-of occurrence—the existence of such phenomena might be considered a very rough diagnostic of the possible existence of dialect leveling.
Previous studies of dialect leveling have, when answering the actuation question, generally focused on dialect leveling as a result of individuals moving into an area, whether due to the creation of entirely new speech communities (e.g., Kerswill 1995, 1996b; Kerswill and Williams 2000) or migration into an already established speech community (e.g., Thomas and Bailey 1992; Berni 1995; Lambert 1995; Herold 1997; Thomas 1997; Boberg and Strassel 2000). In the latter cases, dialect leveling is found when—to generalize—large-scale migration into a region occurs, creating pressure on features of the previously established local language variety. In Waldorf, we have an already established speech community (Waldorf first appeared on maps in the mid-nineteenth century) experiencing changes in rates of /ay/-monophthongization during the twentieth century, but an explanation based on migration into Waldorf does not explain the change.

As can be seen from the rates of /ay/-monophthongization in Figure 3, as well as the VARBRUL weights shown in Figure 7, /ay/-monophthongization was declining well before the middle of the twentieth century. Although the group of individuals born between 1940 and 1949 favor /ay/-monophthongization, they do so less than those born between 1920 and 1939. Similarly, those born between 1920 and 1939 favor /ay/-monophthongization quite strongly, but less than those born before 1920. As can be seen from Figure 2, though, large-scale migration into the Waldorf area did not start until the years following World War II. As a result, an argument based on immigration cannot explain the beginning of the move away from monophthongal /ay/ in the Waldorf speech community. It should be noted, though, that by the end of the twentieth century, Waldorfians certainly had more contact with speakers of other varieties than they did just after World War II. This came about not just because of migration into the area but also because increasing num-
bers of Waldorfians began commuting to Washington, D.C. for work in the intervening years (Edelen et al. 1976, 74). This dialect contact may have accelerated the trend toward diphthongal /ay/, or it may have caused the variation to tend toward a lower level than it might have otherwise, even though it does not appear possible to name it as the cause of the trend.

Another possible explanation is that the changes in Waldorf reported here reflect local cultural changes, along the lines of the possible connection between changes in Southern American culture and changes in Southern American English that Bailey and Tillery (1996, 316) have proposed. This possibility is more difficult to analyze with respect to Waldorf than an argument based on immigration because there is no single statistic one can use to measure culture. However, historians have noted that the Waldorf area’s shift from an agricultural economy to a service economy did not begin until at least two decades after World War II (Johnson and Karpia 1976, 9), and once again the shift in monophthongization of /ay/ predates such a cultural shift.

Another possibility is that the change in Waldorf /ay/-monophthongization is the result of processes completely internal to Waldorf English. That speakers in northern Charles County at one time participated in /ay/-monophthongization to some extent is fairly clear from previous dialectological studies (see Kurath and McDavid 1961, map 47). However, the history of /ay/-monophthongization in the South is fairly complex. As Bailey and Tillery (1996, 313) and Bailey (1997, 266) point out, /ay/-monophthongization in the South is a relatively recent phenomenon, only beginning to appear in, for example, Texas around the turn of the twentieth century (Bailey 1997, 262). Given this, it seems reasonable to suppose that /ay/-monophthongization is also of relatively recent vintage in Waldorf and that its appearance at a relatively high rate among the oldest speakers in this study was the result of a change in progress to that point. If this is the case, and having ruled out contact from outside the Waldorf speech community as the trigger for the reversal of /ay/-monophthongization, it makes sense to posit that the trigger for Waldorf’s trend toward diphthongal /ay/ came from inside the community rather than from outside of it. That is, what occurred in Waldorf in the years preceding World War II may well have been a case of a sort of dialect contact that might be found in any speech community where there is a change in progress—there are forms used differently by individuals in different generations. Although children generally continue changes in progress (Labov 2001, 308), they are still exposed to older models as they acquire language, and they have the option to follow older patterns. If enough children in a particular age cohort were to do this, then a reversal of a change in progress could occur as a result, in a way creating a new change in progress in the direction of older forms. Given the literature that reports on changes in progress, this does not seem to be the usual way that language changes, but it is still a possibility.
If we accept the notion that a speech community experiencing a linguistic change in progress can be considered a dialect contact situation (though perhaps a weak one, one that might better be called “generational contact”), we expect such communities to act in ways similar to communities that are clearly cases of contact between multiple distinct varieties of a language. Kerswill and Williams (2000) provide a list of eight principles that language changes in dialect contact situations generally follow. The first three are important for the discussion here, as they deal with the linguistic outcomes of dialect contact (the following are all from Kerswill and Williams 2000, 84):

1. Majority forms found in the mix, rather than minority forms, win out.
2. Marked regional forms are disfavored.
3. Phonologically simple . . . features are more often adapted than complex ones.

Within a single speech community, it does not seem possible for (1) to hold—the nature of a change in progress that progresses eventually to completion is that a minority form wins out over a majority form. Principle (3) does not seem to hold in the case of /ay/-monophthongization in Waldorf, as the diphthongal form is being re-established in favor of the phonologically simpler monophthongal form. Principle (2), however, merits some attention. In the early twentieth century, monophthongal /ay/ was very restricted regionally in the southern mid-Atlantic. Crucially, Kurath and McDavid (1961, map 47) report that monophthongal /ay/ in "wire" was found consistently in Charles County, but only occasionally in surrounding counties of Maryland and not at all across the Potomac River in Virginia. Therefore, residents of Waldorf would likely have been aware that their tendency toward monophthongal /ay/ was a regionalism, and under principle (2) that tendency would have been ripe for reversal.

This brings up two important issues that require further research. The first is local—the pre-twentieth-century history of /ay/-monophthongization in Waldorf needs to be researched so that /ay/-monophthongization’s beginning and peak there can be documented. The second is more geographically wide-ranging—parallel cases of a local norm being reversed in the absence of any large-scale migration into the speech community need to be found, so that the question of whether Kerswill and Williams’s (2000) principle (2) can generally explain such phenomena can be answered. In addition, such research needs to be done to explain why some regional changes do catch hold despite their regional nature, even to the point that they spread geographically and can no longer be considered regional forms.

In the end, though, this study of the reversal of /ay/-monophthongization in Waldorf requires those of us researching language change to step back and, perhaps, define our terms a bit better. We speak of “dialect contact,” but there are sev-
eral kinds of dialect contact. These range from dialects settled near each other, to
new towns being formed, to a region experiencing massive immigration, to the cur-
rent case (a sort of dialect contact between generations within a single speech com-
munity). If we can actually confirm that all these situations follow the same gener-
alized principles, not only will we be closer to explaining how language change
occurs, but we will be closer to developing a truly predictive theory of language
change.

Notes

1. This result is possible because no interviews for the Linguistic Atlas of the
Middle and South Atlantic States (LAMSAS) were conducted in Waldorf itself.
The LAMSAS interview marked on some maps as having been conducted in Wal-
dorf was actually conducted with a Gallant Green native in Gallant Green, a nearby
but (at that time) very separate community.

2. The “glide present” category was originally coded as “full glide” and “weak
glide,” but that distinction was not found to produce meaningful results.

3. Coding was also done for whether each instance of /ay/ occurred before a
word break, a morpheme boundary within a word, or no such break. However, be-
cause this classification was not independent of all other factors (a word boundary
is very often accompanied by a syllable break, for example), it was left out of the
analysis presented here.

4. Fit to an exponential curve was chosen because this is a change that appears
to be nearing completion, and so it was felt that a curve would better approximate
the closing stages of an S-shaped curve than a straight line would. For complete-
ness, though, the fit of the data to straight lines is $R^2 = 0.9598$ for the men and $R^2 =
0.6413$ for the women.

5. All the differences between categories shown in this chart are significant at
least to a level of $p < .001$, as is the case for the data on the chart as a whole.

6. Once again, all the differences shown in the chart are significantly different
to at least a level of $p < .001$ for each pair and for the data as a whole.

7. The differences listed here are all significantly different to a level of $p < .001$.

8. To conduct this analysis, I used GoldVarb version 2.1 running over ARDI
Executor/Win32 version 2 for Microsoft® Windows® 9x.

9. As is generally done, VARBRUL weights greater than 0.500 were taken as
favoring the variable (in this case, /ay/-monophthongization), and values less than
0.500 were taken as disfavoring it, with distance from 0.500 reflecting the strength
of the favoring/disfavoring effect.

10. I do note that all of this, of course, rests on the assumption that /ay/-
monophthongization is a relatively recent (as in late nineteenth or early twentieth
century) phenomenon in Waldorf; this assertion needs to be verified in future research.

11. *Wire* is referenced here because full monophthongs were not reported in other words.

**References**


Yorkshire, largest of the historic counties of England, so large that it was early divided into three thriddings, or “ridings,” is home to the oldest of the country’s regional dialect societies. Yorkshire speech has been documented often and well, both in the Transactions of the Yorkshire Dialect Society, published since the society’s founding in 1897, and in glossaries compiled in the nineteenth century, as evidenced by the numerous sources used by Joseph Wright for his comprehensive English Dialect Dictionary (1898-1905). Wright himself authored a monograph on the speech of Windhill (near Bradford, west Yorkshire) in 1892. The eighty Publications of the English Dialect Society in the late nineteenth century include twelve glossaries of Yorkshire speech in addition to three reprintings of specimens of Yorkshire dialect literature.¹

As is true of many other English counties, the record of Yorkshire speech began to take shape in earnest only in the early nineteenth century. What little is known from before that time is based mainly on local dialect literature (i.e., representations of Yorkshire speech in fiction and drama), scattered comments by early grammarians and orthoepists, and one early glossary.² Much of this material pertains to northern
English in general rather than to Yorkshire specifically. As scholars seek more fully to reconstruct regional varieties such as Yorkshire English (especially their grammar), they face the challenge of finding and interpreting valid early material. Not only is little known about Yorkshire speech before the nineteenth century, but existing information is almost entirely on lexis, with a small amount on pronunciation. This article addresses the problem of recovering and analyzing colloquial language from an earlier period by a case study of a document written in Knaresborough, west Yorkshire, in the late eighteenth century. Its goals are twofold:

1. to show, by exemplifying and examining selected features, that the language of the Knaresborough document is consistent with other records of Yorkshire speech and to see what the document can add to our understanding of Yorkshire dialect from a period earlier than most other sources;
2. to demonstrate the quantitative use of unselfconscious documents in reconstructing regional varieties of English. These include items written by persons of limited literacy for a private audience, most often family members or close associates. Such documents may lack conventions of punctuation and have inconsistent capitalization, but their language is typically far from erratic. Indeed, quantitative analysis reveals that they show many features consistent with other, more modern studies of regional speech, and others for which information is unavailable from elsewhere.

In recent years, linguists have made strides in reconstructing nineteenth-century regional and social varieties of American English through the quantitative analysis of letters and other written documents from less educated individuals (e.g., Montgomery, Fuller, and DeMarse 1993; Montgomery 1999). This paradigm of research analyzing documents from semiliterate writers has come more slowly to regional varieties of British English. For these, the principal eighteenth- and nineteenth-century sources used by scholars continue to be literary dialect and occasional commentary by observers, grammarians, and the like (Ihalainen 1994). As demonstrated in Giner and Montgomery (1997), letters from English emigrants to the United States often provide important, quantifiable data on certain grammatical and phonological variables. Such colloquial documents can clarify the dialect geography of Britain in the first half of the nineteenth century, one or more generations before collections were undertaken by the English Dialect Society and by Alexander J. Ellis. Though they are not as widespread or as voluminous as emigrant letters, other manuscript documents like the one examined here throw considerable light on regional variation in earlier periods. Indeed, because of their naive character, they can provide linguistic evidence unavailable from other contemporary sources, making their use imperative for the reconstruction of earlier regional varieties.
Of course, manuscript documents present only a partial—and sometimes an uncertain—view of the speech of an earlier day. Each document or set of documents must be evaluated on its own merits and with relation to other contemporary evidence of local speech. The manuscript here discussed is similar to emigrant letters in being a personal, unselfconscious item not written for dissemination or publication, one that reveals much of the writer’s speech. Unusual in many ways, it is a daybook in the form of more-or-less daily reports by the “Master” of the Knaresborough workhouse (i.e., the administrator of an institution giving refuge to town’s paupers) addressed to the “gentlemen,” a committee of men who oversaw its finances and management. Knaresborough today is a small town in north Yorkshire, north of Leeds, near Harrogate, but two hundred years ago it belonged to the West Riding of Yorkshire. Ellis ([1889] 1968) included it in the area corresponding to Variety ia of his Mid Yorkshire District, D30. The nearest locality covered by the Survey of English Dialects (SED 1962) was Y18, Spofforth.

Built in 1737, the Knaresborough Workhouse was by the late 1780s suffering badly from mismanagement, and the committee overseeing it hired a new, salaried master to ensure the solvency of the institution and the welfare of the residents, who consisted of older people, unmarried mothers, and children. William Borrow, the new master, was a man of limited literacy who from August 1791 to September 1792 kept a steady record of the production of each resident and the expenses incurred for their upkeep, in the form of daily summaries and occasional brief narratives of events. He also informed the overseers of the residents’ health, in general and in particular, and any instances of bad behavior. Within his methodical reports, Borrow from time to time recounted interactions with and between residents in which he cited vernacular speech directly. Thus, he produced an extensive and extraordinary document useful not only for dialectological purposes but also for historical pragmatics, local history, and even folklore. A complete analysis of the language in the document is beyond the scope of this article, which will mention some phonological and lexical items and focus attention on selected grammatical features, showing the document’s value for the study of Yorkshire English of two centuries ago.

As will be seen from examples we cite and a more extended excerpt in the appendix, the Knaresborough Daybook has virtually no punctuation. Borrow’s orthography was often inconsistent, with numerous mistakes and naive spellings and with words sometimes spelled in several ways in addition to the standard one. Sometimes he used only nonstandard spellings (e.g., he has fifty-two instances of dresing, but none of dressing). Many of these orthographic forms are grouped in list 1 to indicate the prevalence of nonstandard and phonetic spellings in the document. Most involve silent letters or double consonants, and nearly all are phonetic in one respect or another (especially verb principal parts).
List 1: Orthographic Forms

Forms Involving Silent Letters: abusd, acount (x8), anserd, askt, atempte, atend, backt, before, begd (x3), biterly, bitt (n.), bras (x4), brest, cald (x3) ‘called,’ carrid, chargd, charge, closd, coars, cofin, complyd, confind, conserd, considerd (x2), crid, cry’d, cryd (x2), darkned, dauter, dres (x4), dresers (x3), dresing (x52), drest (x3), endard (x2), endur’d, fals ‘false,’ fitting, fitt, forst (x10), fram’d (x2), gon (x10), handid (x2), hangd, hapned, helpt, hopte ‘hoped,’ hous (x3), imployd (x2), informd (x2), kickt, knitting (x7), las ‘lass’ (x2), laught, liknd, lims, lockt, lookt, plast ‘placed,’ promist (x6), spild (x2), usd (x3), ust (x4), walkt (x2), weid (x2) ‘weighed.’

Other spellings: behaveur, behaveour (x6), burede ‘buried,’ chare ‘chair,’ court ‘court,’ consequence, conserror (x3), constoble, counsil (x3), crismas (x3), cuard ‘cured,’ daughter (x2), deasont ‘decent’ (x2), delegent ‘diligent,’ desember, diserves, docktor (x7), durt (x3), enouf, espeshely, far ‘fare,’ falt (x4), fane (x2) ‘faín,’ fashons, feaver (x2), feavour, fesick ‘physic,’ flower (x2) ‘flour,’ groans ‘groans,’ groser’s, hart (x6) ‘heart,’ harty, hear (x46) ‘here,’ hould, incorag, incouregment, incurigment, indeavers, imuf (x2), isspesaly ‘especially,’ jale ‘jail,’ justis, justise, knowlige, laisy, lickers (x2), metearal, metearal, mutearil, menshon (x7), shous (x2) ‘shoes,’ their ‘there, the,’ winim (x2), winmin (x5).8

Although Borrow’s spellings often reflect pronunciation, the phonological information that can be gleaned from the document is not extensive. The forms in list 1 represent mismatches between spelling and pronunciation that any less literate writer of English might produce. Those in list 2 reveal variable pronunciations and include contextual assimilations, variation in unstressed vowels and syllables, and other phenomena of general colloquial speech or, in some cases, northern or Yorkshire dialect.

List 2: Phonological Forms

aukerd, backah ‘tobacco,’ behant, behint, blenkits, brig ‘bridge,’ bushill (x4), chiris ‘cherish,’ cleathing (x2) ‘clothing,’ cled ‘clad,’ close (x9) ‘clothes,’ creater ‘creature,’ credd (x4) ‘cradle,’ dilatery (x3), divil (x2), divilish, doaf (x1) ‘dough,’ east (x3) ‘yeast,’ eather (x10) ‘either,’ eliven, elsh you ‘else you,’ extraderly, extraudnery, extroaney, fachon ‘fashion,’ famaly (x15), famley, fand (x2) ‘found,’ fatching (x2) ‘fetching,’ furder ‘further,’ gardin (x2), hartley ‘heartily,’ honist, kitchin (x2), Knaisborough (x3), mencing ‘mentioning,’ misforton ‘misfortune,’ modarly, neather (x16) ‘neither,’ nesersery, parcil, parsils, pety (x3) ‘pity,’ pinsh ‘pinch,’ plasher, serveraty, soas (x2) ‘sauce,’ tonetother ‘the one [to] the other,’ tone to tother ‘the one to the other,’ usaal ‘usual,’ ushall, usual (x209), vitels ‘victuals,’ wans (x7) ‘wants,’ whent (x2) ‘went,’ whome (x4) ‘home,’ wich ‘which.’
According to the *Oxford English Dictionary* (OED 1993) and the *English Dialect Dictionary* (EDD 1898-1905), many of these forms represent common colloquial or popular pronunciations in England in the eighteenth century, including *backah*, *blenkits*, *bushill*, *chirish*, *divil*, *eliven*, *famley*, *gardin*, *parcil*, *pety*, and *usaallushallusial*.

Other forms were and still are common in northern England, including *behant*/*behint*, *brig*, *clid*, *creator*, *credle*, and *fand*. The realization of the definite article in *tother* (< OE *ðæt oðer*) is reported by the *English Dialect Grammar* (EDG 1905, §371) for Scotland, Ireland, and England and by the *Survey of English Dialects* (SED 1962) not only in the North but also in parts of the Midlands and the southwest of England (as indicated in Barry 1972, 178-79 and map 5). The final consonant /f/ in *dough* is recorded by the SED §V.6.3 in most Yorkshire localities, as well as in Lancashire. The EDG (1905) indicates that it was a common pronunciation in these counties as well as in Northumberland, Durham, Westmoreland, and some counties of the northwest Midlands. *Eather* and *neather* suggest a pronunciation with /i:/ in the first syllable, still widespread in the North and Midlands of England (Kolb et al. 1979, 129). The SED (1962, §V.6.2) records no instances of the loss of initial /j/ in *east* ‘yeast’ in Yorkshire or other northern counties. However, EDG (1905, §248) shows such a variant pronunciation in south, north, and west Yorkshire, as well as Lincolnshire, Oxfordshire, Sussex, and Dorset. The dropping of initial /j/ is, according to Upton and Widdowson (1996, 37), “a feature of non-standard speech particularly in the North and West Midlands, rather than that in the South-west, where its occurrence is limited to the extreme southwest of Cornwall.”

The dropping of initial /h/ is widespread in northern England today and was also in the mid-nineteenth century (Giner and Montgomery 1997, 172-75). The Daybook author spells *ch* where it is required by etymology (e.g., *humbly*, *hursill* ‘herself’, *husband*), suggesting that initial [h] was retained in Knaresborough in the late eighteenth century. The only possible evidence for its loss is in the form *umer* (which presumably indicates the pronunciation *[jumə(r)]*), not attested by the SED (1962) but recorded by the EDD (1898-1905) for Northumberland, west Yorkshire, west Somerset, and Devon. Ellis ([1889] 1968) and the EDG (1905) stated that initial /h/ in the *wh*-cluster was still used in northern England in the mid-nineteenth century, but the Daybook offers little evidence to support this view. Variant spellings such as *whent* and *wich* suggest its writer lacked initial aspiration and used what is the modern-day standard pronunciation in England /hw/. Likewise, the document has little evidence for the loss of postvocalic /r/. The only forms suggesting the pronunciation are *Knaisborough* (used three times), *extraderly/extraudnery* *extroaney* (variant spellings of *extraordinary*), the ‘their,’ *nesersery*, and *serveraty*; the last two are hypercorrect forms suggesting the loss of /r/.

Two pronunciations for *clothes* ([klaʊz] and [kluɔz]) are attested by the SED (1962) in Yorkshire. There is a clearly defined isogloss running from the northwest
to the southeast, with [ɪə] in the northeast and [ʊə] in the southwest (Kolb 1966, 209), with Knaresborough in a transition area. The Daybook form close represents a popular pronunciation of clothes. SED §VI.14.19 records one instance of cleathing in Helmsley, twenty miles northeast of Knaresborough.

Sometimes pronunciation radically affects the form of a grammatical pattern, as in the reduction of have to a (sentences 1-2) or further to zero (3):

(1) I wod fane a had him to a spun a little
(2) I should a holden to it all day
(3) I should ø said more but she is going out to Justis first thing then house

Lexis and Phraseology

The document is particularly rich in dialect lexis, containing numerous forms from Yorkshire and northern England, as shown by examples 4 through 10.

(4) “at cotton miln”
   miln ‘a mill, a factory’ EDD Scotland, Yorkshire, Lancashire, Derbyshire, Nottinghamshire, Lincolnshire."

(5) “they bide a deal of following”
   bide ‘need, require’ EDD North Country, Cumberland, Westmorland, west Yorkshire (EDD states that the sense applies only to things, but this is contradicted by its citation “T’owd gent al bide a good deal o’ waating on”).

(6) “they bide a deal of following”
   follow ‘to look after, attend to, mind’ EDD Yorkshire, west Yorkshire “Too big a house for me to follow.”

(7) “Elizeh Clemshay talking dutch to them at wod give ear to her”
   talk Dutch ‘speak unintelligibly’ EDD s.v. Dutch ‘unintelligible, difficult to comprehend, also used adverbially’ west Yorkshire 3 “The native was foiled in his attempt to understand the gentleman who spoke Dutch,” west Yorkshire 5 “Naa barn, thah tawaks Dutch, ah can’t undrstand thuh.” (The OED has no quotations of dutch with the meaning ‘unintelligible.’)

(8) “Elizh Clemshay walking about teling a dutch tale”
   tell a Dutch tale ‘talk incomprehensibly’ (This expression is not found in the EDD [1898-1905], OED [1993], Dictionary of American Regional English [DARE 1985-], Historical Dictionary of American Slang [HDAS 1995-], etc.)

(9) “Ann palferman Nutty and my wife is as throng as three in a bed”; “Elizeh Clemshay washing her own Close very throng with that”
   throng ‘crowded, very fully attended or frequented, busy.’ EDD ‘busy, occupied, absorbed in work; of work: pressing’ Scotland,
North Country, Northumberland, Durham, Cumberland, Westmoreland, Yorkshire, Lancashire, Cheshire, Derbyshire, Nottinghamshire, Lincolnshire, Leicestershire, Hampshire. *OED*
Now Sc. and north. dial. (The *OED* has more than 300 quotations with *throng*, most with the meanings ‘crowd, crowded,’ but only one for “as throng as three in a bed,” from 1770 *Lett. Jas. Murray, Loyalist* [1901] 134.)

(10) “hear is Hinderson poorly a poor weak Caleley Creater”: *caleley* ‘in poor health,’ *EDD* s.v. *cail* Yks. ‘To be in poor health, to suffer.’ west Yorkshire “Miss Julie was always cayling.” Hence *Cailing* ppl. adj. ‘ailing, weakly’ west Yorkshire, Yorkshire; *Yorkshire N&Q* (1888) II. 109. west Yorkshire 3. (No similar variant recorded in *OED*. The term might be related to *quailing* ‘diminishing, becoming weak; losing hope or courage, etc.’)

Lexical variation in one case deserves special note. The Daybook author uses two regional plural forms for children—that is, *childer* (x200+) and *barns* (x7), the latter apparently to express a degree of intimacy. *Children* occurs once.

**Grammar**

It is the grammar of the Knaresborough Daybook that deserves greatest attention. We will note, and in some cases quantify, a number of features. The document exhibits many patterns of general colloquial speech such as double negatives and the historical present, as well as others of more limited distribution, such as modal *mun* ‘must’ (sentence 11), a form found in northern England (Borrow also writes *must* as *mut* and *mud*).

(11) I both mun have it done and will have it done

The occasional narrative vignettes in which Borrow recounts direct speech display a variety of emotions not found in the rest of the Daybook record. In these brief sections are found frequent use of *thou* (probably pronounced [ðu] in the local speech;12 see *EDG* 1905) and *thee*. Within the document, variation between *thou* and *you* was obviously complex, but it generally followed a pattern lost in the literary language two centuries earlier (but still prevalent in regional dialects a hundred years ago, as noted in *EDG* §272). The choice of a *th-* form or a *y-* form depended on the status and relationship (both of which could also be manipulated by the choice of pronoun) between the speaker and the addressee, as well as emotional intensity. In the vignettes, both the master and residents use *th-* forms in more emotive contexts to express, for example, contempt in (12) or endearment in (13).
(12) Mary Outhwith Come from Miln Viantly Stand back she said or I will make you drives all of fire Nutty was set with Child of her kne saith Mary Stand off with the thou bitch and whore or I will make the saith Nutty If I be a whore thou art a thif of that they tuke to blows

(13) she said I am sick and pain in my head I said I will get the any thing aples pears gins bread wine or any thing thou desires she lay still

A number of grammatical features occur with variation sufficient to quantify. For instance, for the relative marker in restrictive relative clauses, that (x59) is dominant. However, the northern form at (either derived from Norse or a reduction of that) also occurs with some frequency: twenty-two times in restrictive clauses (see 14 and 15) and once in a nonrestrictive clause (16). Zero occurs seven times, all in existential clauses (as in 17). Relative who and what never occur in either restrictive or nonrestrictive clauses.

(14) if I have nothing at is good I have nothing that is bad
(15) that at Cannot be Quard [i.e., cured] must be induerd
(16) mary thorp at is bad of sight
(17) their is mary palferman ø Spins Six hanks a day

As with many varieties of English, the paradigm of copula/auxiliary be shows variation here, as seen in Table 1. In the first person singular, I is reflects a pattern that is both historic, noted for northern Middle English by Wright and Wright (1926, 202), and contemporary, documented by the SED for five northern counties, including Yorkshire (Upton, Parry, and Widdowson 1994, 494). The variation here is probably best analyzed as externally based (i.e., competition between I is and the standard, literate alternative I am). There is no clear contextual conditioning, although I is is the more frequent form as the auxiliary (13/23 of the I is tokens precede the progressive or past participle, as in 18-19, whereas only 10/13 tokens of I am do). How-

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present Tense of be</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>First person</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Second person</td>
</tr>
<tr>
<td>Third person</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

NOTE: be occurs in the subjunctive, as in he Certainly will be hang if he be on land; “NP” refers collectively to all nominals other than personal pronouns.
ever, that both forms can take the same adjective (afraid, seen in 21-22), suggests no obvious semantic or syntactic complementarity between the two.

(18) I is now going to die
(19) I is forst to use my rod
(20) I is very ill set with them
(21) I is afraid their will be Sum falt fun
(22) I am afraid you will think bills will be hard

In the third-person plural, the variation between is and are depends very much on the type of subject. We see here the operation of the familiar verbal concord rule for the northern half of Britain, cited as long ago as by James Murray (1873) in his *Dialect of the Southern Counties of Scotland*. Ihalainen (1994, 221) calls this the “northern subject rule,” according to which, “plural present-tense verbs take -s, unless they are immediately preceded by a personal-pronoun subject, as in *They peel them and boils them* and *Birds sings.*” A speaker or writer who follows this historical rule would mark with suffixal -s a third-plural verb having any subject other than an adjacent they. A verb having a first- or second-person pronoun subject would also take -s only if not adjacent to its subject.

In Giner and Montgomery (1997), we demonstrated the vigor of this rule for both lexical verbs and copula/auxiliary be in a set of mid-nineteenth-century emigrant letters from Huddersfield, west Yorkshire. This pattern is also noted in such modern-day studies of Yorkshire English as Hedevind (1967), Petyt (1985, 237), and the *SED* (1962). Table 2 details the patterning of verbal concord in the Knaresborough Daybook for two types of verbs in the present tense.

However, this table is actually misleading because for all three apparent exceptions, the verb is not adjacent to the subject (23-25).

(23) they bide sum driving and is dilatery
(24) they do their work bides a little driving
(25) they frump him and wants to have all bras they can get for their self

<table>
<thead>
<tr>
<th>Form</th>
<th>Type</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>copula/auxiliary <em>be</em></td>
<td>they are</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>they is</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>NP is</td>
<td>44</td>
</tr>
<tr>
<td>lexical verbs (for -s)</td>
<td><em>They</em> + Ø</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>they + -s</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>NP + -s</td>
<td>48</td>
</tr>
</tbody>
</table>
For verbal morphology, the Knaresborough Daybook presents one puzzle. *Saith* (*x*26), *doth* (*x*36), and *hath* (*x*51) occur throughout the document, in both reported speech and elsewhere, nearly to the exclusion of the parallel (and expected) forms *says* (*x*1), *does* (*x*0), and *has* (*x*4). The forms in *-th* are unattested in the speech of Yorkshire or northern Britain of any period, so far as we can determine. They are known only from southern England, where they occurred in the third-singular and throughout the plural well into the sixteenth century in speech at least variably (Stein 1987). Jespersen (1961, 19-20) states that in the first half of the 17th c. *s* must have been the ending universally in ordinary conversation, and we have evidence that it was even usual to read *s* where the book had *th*, for Richard Hodges (1643) gives in his list of words pronounced alike though spelt differently among others *boughs boweth bowze; clause claweth claws*.

The suffix continued as an orthographic variant into the eighteenth century, and in certain registers (e.g., prayers and other religious discourse) it continues as a self-conscious archaism. The forms in the Knaresborough Daybook might appear to be simple intrusions from the Authorized Version of the Bible (1611) or other prestigious prose, except that no other artificial orthographic forms show up in the document and except for sentences like the following:

1. their Childer *hath* done their work as Usial today
2. poor people *hath* been very much out of health
3. but when fair words *doth* not do I is forst to use my rod
4. they have done their work very well and *hath* pleasd the owner of the worsit very well
5. they five Spins Twenty four a day and *doth* their work

In the Daybook’s plural contexts, forms in *-th* conform categorically to the northern concord rule, which was a speech-based pattern allowing the inflection only on a third-plural verb having a subject that is a noun (26-28) or a nonadjacent pronoun (29-30). Apparently, the Daybook author treated *-th* as an orthographic variant of *-s* and generalized it to contexts in which his speech had *-s*. This would certainly attest to the strength of the northern concord rule in eighteenth-century Yorkshire, but it does not account for all the data. Why does *-th* occur only on *doth, hath, and saith* and not on other verbs (other than one instance of *eath*, in “she sath she eath nothing”)? If the two endings represent orthographic variants in only certain words, does this reflect a regional pattern?

Finally, a few observations about the occurrence of zero copula/auxiliary in the Daybook are warranted. The possibility of this feature in Yorkshire English was first brought to our attention by Melchers (1972, I:138), who notes that “the present
progressive is often formed without an auxiliary” in modern-day west Yorkshire English not far from Knaresborough. The evidence from the Daybook is striking and quite possibly indicates a heretofore-unknown feature of eighteenth-century colloquial British English, though it comes from only one individual and may in part be idiosyncratic. Since zero copula has apparently been unnoted for any other variety of British English, it is important to ask what would constitute sufficient evidence in a historical document to posit the existence of this feature in the speech of an earlier day. Excerpt (31) contains six zero forms, four before such present participles as doing and spinning, and no overt forms of the copula/auxiliary verb. Zero forms like these, typical of list or telegraphic style, are extremely common in the Daybook when the author cites workhouse inmates and their daily activities. They represent a commonplace stylistic option and are not of interest to us.

(31) ann palferman mary Outhwith ø doing house busness Elizh Clemshay doth not Care for no work at all Wm Clemshay ø at his work as Usial Sarah Eyeley Jane prince Ellin Robinson Sarah Forist ø Spining line Richd Garman Richd mason as Wm Brown ø teaching Childer Andrew Steel Elizh Thorp Hannah Dixon ø winding Jane Whelas Jane Foster ø as Usial

On the other hand, thirteen examples (32-44) occur outside such passages, two of these in reported speech. Zero forms occur in both singular and plural contexts and follow both vowels and consonants, most often t (35-42). They precede nominals, adjectivals, and other elements (but a progressive only once, in 44) and express a range of temporal states from punctual (38) to general truth (33).

(32) all hands ø at work at is fit to do anything
(33) Gentlemen their is not a deal to Steal on when they ø all served and when you have all Considerd it over
(34) she ø nothing but skin and bone and a boundance of Cumber
(35) Now Gentlemen hear is two out of the house that ø Gibson two boys and the rest I have to order their work
(36) Gentlemen a good trusty servant that ø trusty to his master he is sure to be good to him self
(37) ye all want to be kept for nothing doing like drones that ø in a hive
(38) mary I said what ø the matter she said they Call me thief
(39) their will be Sum falt fun [= found] with the grosers bill that it will be larger then what ø Usial but the Acasion is poor people hath been very much out of health
(40) pal what ø the matter with the [= thee]
(41) a woman that hath a Child Sucking on shud have Sum thing that ø good and Nurishing
(42) Seven Old wimmin that ø not worth Sixpence per week
(43) now I said I see who ø Master but no matters blood shed—
This evidence is quite provocative, given the lack of documentation of zero copula in British English, but is hardly sufficient to suggest what, if any, phonological, grammatical, or semantic factors might condition the feature. It is difficult to state how prominent zero copula was two centuries ago, but the Daybook provides cases that are unambiguous and cannot be attributed to telegraphic or list style.¹⁵ It shows no apparent similarity to zero copula in varieties of American English such as African American Vernacular English, and since colonial-era emigration from Yorkshire to North America was sparser than from many other parts of Britain, a trans-Atlantic connection is doubtful. However, as the first substantial body of evidence for zero copula from British regional English, it requires much further scrutiny.

Discussion and Conclusion

In this article, we have presented some linguistic features of a remarkable document written in west Yorkshire in the late eighteenth century by an individual whose limited literacy reveals many traces of his speech. Nearly all of the language in the Knaresborough Daybook can be localized to Yorkshire with the aid of sources such as the English Dialect Dictionary. The manuscript itself gives an earlier view of many aspects of the lexis and grammar of Yorkshire English than is currently on record.¹⁶ Its evidence for variation in verbal morphology is especially noteworthy. The northern British character of the language is argued by the prevalence of I is, the northern concord rule, and other features. The document suggests that zero copula also occurred in some parts of northern England two centuries ago. In sum, this article confirms the wider and quantitative usefulness of unselfconscious documents to reconstruct earlier regional varieties of British English. Furthermore, it extends our knowledge of earlier Yorkshire English and provides the most substantial body of evidence on eighteenth-century Yorkshire English now known.

APPENDIX

Excerpt from the Knaresborough Daybook

Jeny 14 [1792] Gentlemen yesterday 12 inst. we had an
Ingadement betwixt Nutty and Mary
Outhwith Mary Outhwith Come from Miln
Viantly Stand back she said or I will
make you drives all of fire Nutty was
set with Child of her kne saith Mary
Stand of with the thou bitch and whore
or I will make the saith Nutty If I be a whore thou art a thif of that they tuke to blows my wife see the frays Calls out Come Come they fighting I heard them and went what hear to do but mary wod fight I put both into our house and their I had a fair hen battle but Nutty downd mary presantly and proved therby Concerer now I said I see who Master but no matters blood shed Jeny 16 Monday Gentlemen all the wimmin in the house is frameing 100 Shifts and Sowing them at that Can sow for we Expect to have manty maker and taylor both which will be a great Expense this month for Mr. Allinson saith he will put all in deasant repare this month so I told Them they mud be all good and Carefull and not ware all their bras of halfpeny Cakes and peny Cakes I said then if every honist body would but sit down and Count what vast sums Small trifels would amount one peny of the day you may think it Clear it will make L1 =3D 10s =3D 4e every year Augh they said we never pinsh belley back will trust but belley will not trust long Gentlemen we have all the paupers at work ats fit to do any thing and to make a menshon of all their names it would stand for nothing Gentlemen this Cleathing is very Chargable but one thing I have to say he or she is the best cooks that can set out a hansan table with a few vitels

Notes


2. Best (1641) is one of the earliest sources for Yorkshire dialect. See Widdowson (1992-1993) for a thorough account of sources for the study of the dialect of Sheffield, West Yorkshire between 1829 and 1984.
3. This is not to discount such scholarship as Davis (1989) on the Paston letters and Nevalainen and Raumolin-Brunberg (1996) on sixteenth- and seventeenth-century correspondence, but these and similar works analyze unselfconscious documents from more literate writers and are mainly concerned with processes of standardization and linguistic change rather than the description of vernacular and regional varieties of language.

4. Strong (1992, 10) suggests that these men, who also formed the vestry of the local church, had asked that a “written record of the day to day events in the house should be kept” to prevent the abuses that had taken place in the past. An overseer of the poor controlled the income generated by the activities of the poor: working at the mill, spinning, and winding. This income funded the workhouse and was used to lessen the poor rate (Strong 1992, 10ff).

5. The first entry of the Daybook was for May 1, 1788, but for more than three years the document contains only accounts and brief notices. Only in August 1791 did entries begin a narrative account of daily activities. The Daybook also contains one entry—for February 5—for after September 26, 1792.

6. An annotated edition of the document is being prepared by the authors.

7. Forms occur once unless otherwise indicated.


9. These spellings may represent the retention of a different vowel, that is, [ɛ] or a diphthong [ɛɔ].

10. Humor without aspiration is well-known in the United States, where general /h/-dropping is not found.

11. Miln was the earlier form. The final consonant was lost in Middle English (as it was in kiln) but was retained in English regional dialects.

12. As is well documented in the Survey of English Dialects (SED 1962) and elsewhere, the Great Vowel Shift did not carry through to completion for Middle English vowels /i/ and /u/.

13. A similar point is made by Percy (1991), who analyzes variation between third-singular -s and -th in Captain James Cook’s “Endeavour” Journal kept from 1768 to 1771. Cook was born in Yorkshire in 1728 and had only a year or two of formal education in Ayton. Percy finds that Cook used -th on many verbs, especially on hath (x80) versus has (x33) and doth (x20) versus does (x4).

14. Derek Britton (personal communication, 2001) advises us that -th in these nonadjacent plural contexts “are practically the norm among 16th century northern scribes. It’s an orthographic type of hypercorrection. They know that the formal written variant of 3sg has -th, which they duly replicate. But since [the concord rule] in the non-adjacency condition requires pl form be identical with the 3sg, they generalise the -th 3sg to pl, while in speech having {s} for both, one imagines.”
15. Martin and Tagliamonte (1999) have recently documented zero copula in the contemporary speech of Wheatley Hill, County Durham, and the structural constraints on its occurrence. They also found no evidence that it correlated with semantic features. Shorrock has found several instances of zero copula or auxiliary in his forthcoming edition of Mason’s *Glossary of the Dialect of Longtown, Cumbria*.

16. Another notable linguistic analysis of colloquial Yorkshire English of the eighteenth century is MacMahon (1980).

References


Reviewed by Patricia Cukor-Avila
University of North Texas

Every spoken language is dynamic and thus in a continual process of change. McWhorter introduces readers to this fundamental principal of linguistics in the first paragraph of Spreading the Word, where he sets the stage for the discussion of variation in American English and linguistic equality in light of society’s beliefs about “good” and “bad” language, two of the main themes explored throughout the text. Its short length (76 pages of text) and informal style make this book accessible for an audience unfamiliar with linguistic concepts yet interested in issues of standard versus nonstandard language, specifically in educational settings.

Spreading the Word consists of a brief introduction, four main chapters—each followed by short exercises—and ends with suggestions for further reading. Chapter 1, “I Hear So Much Bad Grammar These Days,” begins with the classic comparison of the Lord’s Prayer written in Old and Middle English to demonstrate variation and change over time in English. This point is further illustrated with a short discussion of how Latin was transformed into the five Romance languages. The remainder of the chapter addresses dialect diversity in the United States in terms of what McWhorter outlines as the six “simple points” of dialects (6-15):

1. Any dialect you hear, no matter who is speaking it or what they happen to be saying, is the product of the same kind of change that turned Latin into French.
2. Colloquial dialects develop alongside standard varieties, not from them.
3. Standard dialects are chosen according to geopolitical accident, not according to anything inherent to the dialect itself.
4. Colloquial dialects are as complex and nuanced as standard ones.
5. Our natural sense that one dialect is “better” is based purely on sociological associations, not anything about the dialect itself.
6. If a person speaks a colloquial dialect, it is not a symptom of an inability to speak a standard one.
Each of these points is summarized using examples from English and other languages (such as German, Bulgarian, and French), illustrating quite effectively to the novice reader that dialectal variation occurs in all languages of the world. It is in this section (specifically point #4) that McWhorter first discusses dialectal variation in present-day English by introducing readers to the form and function of habitual invariant *be* in “Black English” (McWhorter has chosen to use the term *Black English* rather than African American Vernacular English [AAVE] or African American English). The chapter concludes with advice for educators that “there is no such thing as a native speaker of English speaking bad grammar” (15)—hence the argument that *goed* for ‘went’ is developmental and that multiple negatives or invariant *be* are “articulate constructions in *other kinds of* English that happen not to have been chosen as standard” (15). McWhorter reminds readers that “the job of school is to *add a new layer* to a child’s speech repertoire, *not to undo* the one they already have” (15).

Chapter 2, “It’s Just Slang, Isn’t It?” addresses the differences between accent, dialect, and slang and the misconceptions about them that often have direct implications in classroom settings. Each of these terms is discussed, again, with examples from several languages (including English). McWhorter’s overall message in this chapter is that linguistic complexity is unrelated to “rich” communication and that “there is nothing inherently ‘second-drawer’ about colloquial English dialects, especially because they contain many hidden complexities of their own” (32). He draws on two examples from AAVE—the use of *had*+ past for simple past and intonation patterns—to argue this point. It is here that McWhorter’s simplified versions of complex linguistic processes, referred to as “little mismatches between standard English and Black English” (33), become problematic for researchers in the field. A brief explanation of how the pluperfect is used in Standard English precedes a personal account of the author’s confusion when his cousins “would pile *had* upon *had* without ever settling back into the regular past” (33). This is followed by a one-sentence example demonstrating that “related speech varieties differ in how they use their *have’s*” (34): French *J’ai vu* for “I saw” but Spanish *He visto* for “I have seen.” Rather than support the notion that dialectal varieties of English (in this case AAVE) are systematic and rule governed, this example does nothing more than reinforce the standard versus nonstandard dichotomy that McWhorter is trying to avoid throughout the book. Unfortunately, this dichotomy is replayed in the exercises that follow where it is suggested that readers review and compare passages written in “Standard” and “Black English” and then translate from one dialect to the other.

Chapter 3, “They Just Mix Them Up!” addresses two main topics: code switching and creole languages. The description of code switching focuses on switching between Spanish and English (although there is also an example of French/Wolof switching for comparative purposes), which is very relevant for teachers and ad-
ministrators in schools with Spanish-speaking immigrant populations. McWhorter emphasizes that code switching is not an “impairment” or “a symptom of someone having never learned to truly speak a language” (42)—both of these points are essential information for readers who are unfamiliar with the linguistic aspects of code switching. The following section includes a brief discussion of lexical change resulting from language contact that serves as the introduction to the discussion of creoles taken up in the second half of the chapter. McWhorter’s overall goal in describing creole languages is to dispel the myth that they are not “real languages” (50). He does so quite effectively through an example of the grammatical complexity of the use of “to be” in Haitian. Readers are reminded that “creole languages are further demonstrations of what we have already seen, that there is no such thing as fluent human speech that is bad grammar” (52). The chapter ends with short histories and sample passages of three creoles in the United States: Gullah, Louisiana Creole French, and Hawaiian Pidgin.

McWhorter concludes Spreading the Word with a chapter titled “The Linguistic Rain Forest,” this being a metaphorical reference to the major presence of languages other than English “that now occupy center stage” (62) in the United States today. The chapter summarizes the linguistic histories (briefly with examples) of eight such languages: Russian, Chinese, Vietnamese, Tagalog, Hebrew, Amharic, Hindi, and Korean. McWhorter describes (and compares) each of these languages in terms of its unique linguistic characteristics: morphology in Russian, Tagalog, and Amharic; tones in Chinese and Vietnamese; tense in Hebrew; and semantic and pragmatic distinctions in Hindi and Korean. The discussion in Chapter 4 departs from the standard/nonstandard issue raised in the previous chapters; however, the message to educators is clear: (1) linguistic variation is the norm, not the exception; (2) the linguistic structure of English is by no means the template for languages of the world; and, most important, (3) all languages are linguistic equals.

Spreading the Word is a suitable supplementary text for an introductory undergraduate linguistics class or for a class concerning varieties of English in the United States. The discussion of language ideology and linguistic discrimination in educational settings is much more limited in scope than that in Lippi-Green’s (1997) English with an Accent or Burling’s (1973) work on AAVE for educators, English in Black and White. Readers unfamiliar with linguistics will appreciate McWhorter’s brevity and informal style, although at times the examples are overly simplistic and some may tire of the pop culture humor interspersed throughout the book. For example, the reason why the dialects of Oxford and Cambridge became known as the most prestigious in Great Britain is explained by the following:

This choice was as arbitrary as the fact that in the early 1980s, the torn sweatshirt that Jennifer Beals wore in the movie Flashdance became a fashion craze because Beals was lovely, the movie was a hit, and its soundtrack was
coming out of our bathroom faucets for a year. If you think about it, what’s so
great about a torn sweatshirt? If Roseanne Barr had worn one for a whole sea-
son of Roseanne, not a single woman in America would have taken out scis-
sors to rip up her nice sweatshirts; what made this make sense in 1983 was a
random confluence of glamour and sensory overload. It was the same with
the East Midlands dialect that became standard English. (8-9)

Nevertheless, *Spreading the Word* is a welcome addition to the available materials
for students and educational professionals new to the study of language and dialect
diversity.

**References**

Winston.


Reviewed by Ellen Johnson
Berry College

I have been known to complain that those of us who study language variation should stop arguing about methodology and just get to work. A great deal has been written about the methods of sociolinguistics, largely because when different researchers reach different conclusions, we suspect that the difference lies in the methods used to collect and analyze data. In a field where one of our largest professional meetings is called the “New Ways of Analyzing Variation” conference and a major international meeting is known simply as the “Methods” conference (John-son 1995), do we really need another book on methods?

The answer is yes, and that need is admirably filled by Barbara Johnstone’s work, Qualitative Methods in Sociolinguistics.

Many of us are familiar with Johnstone’s research. Her book on The Linguistic Individual has certainly shaped my own thinking about the causes of language variation or, as she puts it, the reasons for the linguistic choices people make. She is a knowledgeable, thoughtful, and experienced researcher who has the ability to communicate her ideas clearly, which is why Qualitative Methods is so successful. It will be useful both as a textbook and as a guide to researchers who want to expand and refine their methodology.

Finding the right textbooks for a course on research/field methods in linguistics can be challenging. There are some good books available on quantitative methods in dialectology (e.g., Kretzschmar and Schneider 1996; Davis 1990; Preston 1993) and on ways of researching language acquisition (e.g., Nunan 1992), and some cover subfields of qualitative research from the perspective of other disciplines: how to write up field notes (Emerson, Fretz, and Shaw 1995), how to interview (Ives 1980), and how to write ethnography for publication (Van Maanen 1988). However, any social scientist wishing to adequately survey current methodologies must be aware of qualitative techniques and the postmodern critiques of positivism that have led to their popularity. The last time I taught research methods, I sought in vain for a general text on qualitative methods that used linguistic analyses as examples, which is exactly what Johnstone has produced here. There are boxes that summarize and highlight particular topics, and the author also provides some useful discussion questions and assignments that ask students to apply what they are learning.
She addresses the important philosophical questions, such as the following:

What makes a research question worthwhile?
Can we discover truths about the world through systematic observation?
Is deception or covert observation ever ethical?
How can qualitative analysis be rigorous?
How can qualitative and quantitative research be combined?
How can researchers avoid overgeneralization?

And she also addresses important practical problems:

how to ask informants to give informed consent,
how much to tell informants about the purpose of the research,
how to arrange the tape recorder and the microphone,
how to represent speech in transcriptions for analysis and for publication,
what wording and format is appropriate for the resulting journal article.

An important question for me, as alluded to above, was “How have linguists used qualitative research?” Johnstone aptly illustrates her discussions with descriptions of actual sociolinguistic studies that have successfully used these techniques (e.g., Robin Lakoff on introspection, Shirley Brice Heath on participant observation, a consent form used by Mary Buchholz). Wherever possible, she also includes suggestions from the researchers themselves about methods they have used or studied (e.g., Deborah Tannen on the use of “playback,” Dennis Preston on transcription). The chapter on “Standards of Evidence” could have benefited from more examples, but the book as a whole is readable and understandable due to the judicious use of examples that sociolinguists can relate to.

The book does not focus predominantly on the effects that the research setting and the personal characteristics of the field-worker have on the speech that is observed, as do most treatments of methods for studying language variation (cf. papers from a panel on “Fieldwork for the New Century”; e.g., Hazen 2000; Fuller 2000). Johnstone does not gloss over these effects, but this is only one of many topics she covers, and she discusses them in some new ways. For example, at various points throughout the book, she mentions that different personality types are better suited for different tasks, different methods, different settings, and different interests and that researchers should consider what they are good at and what they enjoy in planning their projects. In this spirit, she considers the question of whether insiders or outsiders make better field-workers. She recounts a project where she teamed up with an assistant, Judith Bean, a native Texan who could establish rapport with the informants, after which she (Johnstone) would play her role as an outsider:
Because I was assumed, like all Yankees, to know very little about Texas (even after 10 years there), I could ask dumb-sounding questions whose answers Texans might assume were obvious to other Texans.

This excerpt shows how Johnstone’s sense of humor makes what otherwise could be a very dry subject a delight to read.

Perhaps the relative lack of attention to the effects of the setting and interviewer is due to Johnstone’s attitude toward what kind of data are useful. Rather than succumbing to the limiting ideas that the only speech worth studying is the most casual style of speech and the only interesting groups to study are those whose speech is farthest removed from the standard dialect, Johnstone takes the broader view that any language use can potentially be of value to our field. I have argued elsewhere that this is exactly how we need to expand our horizons, and the call has been taken up by other sociolinguists as well (Johnson 2000; Schilling-Estes 2000). Research that depends on field-workers’ ability to collect samples of the “vernacular” requires a great deal of attention to getting the style of interaction and the setting just right, but that is not the central concern here.

Students often have a difficult time framing language-related research questions from the wealth of ethnographic data they collect. Johnstone’s chapter “Thinking about Methodology,” in which she discusses research questions, is quite insightful, going beyond expected sorts of advice, like choosing a question that you think you can answer using the resources available to you, to candidly considering the pros and cons of students choosing a project because it is related to what their major professor is doing, of whether it is wise to take a radically new approach to a topic (one could become a “pioneer” or one could be marginalized), and of whether to consider the likelihood of grant funding in choosing a topic. She recognizes that research topics may be dictated as much by the demands of securing tenure as by more noble concerns, and this candor is refreshing. I thought the use of the terms deductive and inductive would have been helpful in this section, as she discusses the concepts (i.e., planning the study with a theory in mind vs. letting the work be “data-driven”) without applying the usual terminology.

“Thinking about Methodology” and the following chapter, “Some Legal and Ethical Issues,” both contain a wealth of information packed into a small space (the entire book is only 143 pages). The section on informed consent and the reasons for it is clear and persuasive. It brings to mind the ethics statement of the American Anthropological Association (1998), which says that we should always be open about our purposes with the people we are studying. Johnstone also broaches the topic of using research assistants, and she refers profitably to Cameron et al.’s (1992) work on issues of power in research.
The four chapters that make up the second half of the book are entitled “Thinking,” “Looking,” “Reading and Listening,” and “Writing.” Together they guide researchers through the research process by treating hypothesis generation, observation, analysis, and reporting results, respectively. Here the author includes everything from a philosophical discussion of how we create groups by the way we label them, to Hymes’s classic heuristics for the kinds of data we need, to an excellent section on transcription practices that presents the same data in different ways to show the impact of the choices we make in how to represent speech.

Johnstone notes that the best projects use a variety of methods and perspectives to try to get at the truth about language use, one reason many studies are both qualitative and quantitative. She shows readers some of the advantages of qualitative research while advocating the use of both types of methods; thus, she takes a conciliatory rather than a defensive stance toward the question of which is better. The main criticism quantitative researchers have had toward qualitative methods is the problem of generalizability of results. While statistical tests can tell us reliably about the probability of a phenomenon occurring beyond our sample, in the population at large, qualitative researchers often leave it to their readers to assess whether the patterns found in their studies would hold true in other situations. Johnstone does not discuss this as a problem in as much depth as it warrants, but she does warn researchers not to make general statements, rather to always remember that what they found is true for one particular set of speakers at a particular time and place. Qualitative researchers can mitigate the problem of generalizability by describing their subjects and their settings as thoroughly as possible to help other researchers decide for themselves whether they are like other seemingly similar groups and places.

Linguistics lagged behind other social sciences in adopting quantitative methods. Nowadays, critiques of “objectivity” have led to a great deal of interest in qualitative methods in many fields. Linguists can benefit from these methods and indeed have already perfected many of them, albeit without explicitly explaining how and why they do what they do. We should not lag behind this time. I heartily recommend this book as a guide. Researchers always need to be informed about good research practices.

To sum up Johnstone’s advice, I would say collect plenty of data, become intimately familiar with them, and always think carefully about what you are doing and what your assumptions are. Over and over, she characterizes good, sound qualitative research in sociolinguistics as “systematic observation,” and she provides the reader with a rich body of information about how to be observant and how to be systematic about it.
References


Reviewed by Laurel J. Brinton
University of British Columbia

Pathways of Change: Grammaticalization in English (henceforth PCGE) is a collection of papers originally presented at the Thirteenth International Conference on Historical Linguistics, held in Düsseldorf in 1997, supplemented with a number of solicited papers. The purpose of the volume is “to broaden the range of empirical cases of grammaticalization in . . . English, and thereby cast more light on a number of current themes in grammaticalization” (1). As Olga Fischer and Anette Rosenbach point out in their introduction to PCGE, English, with its dearth of inflectional morphology, would not seem to be particularly well suited to the study of grammaticalization, the historical process that, according to the best-known definition, “consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status” (Kurylowicz 1965, 69). However, as Fischer and Rosenbach also point out, the shift from synthetic to analytic structures in the history of English necessitates grammaticalizations of a somewhat different sort, and the well-recorded history of English provides opportunities for careful case studies, such as those of lovely, pray/prithee, methinks, to, (al)though, and OE onginnan/beginnan and soplice/witodlice found here. The “themes” of grammaticalization studies that figure most prominently in this volume are the directionality of grammaticalization, the distinction between grammaticalization and lexicalization, and the relation between synchronic variation and grammaticalization. My overall impression of this volume is that it presents a variety of papers relating to diachronic change in the history of English, which of themselves are worthwhile and very interesting, though not all of the papers are as tightly focused on grammaticalization as the title of the volume—and the informative introduction to grammaticalization provided by Fischer and Rosenbach—might lead one to expect. Nonetheless, the volume makes a significant contribution to the study of grammaticalization.

Most germane to the topic of grammaticalization are eight of the papers in PCGE, seven of which are empirical studies that concomitantly raise larger issues concerning grammaticalization and one of which is primarily theoretical. In conjunction with a close examination of methinks, Ilse Wischer (“Grammaticalization
versus Lexicalization: *Methinks There Is Some Confusion*) interrogates the vexed distinction between grammaticalization and lexicalization, concluding that they “are not at all contradictory processes . . . but operate on different levels of the language” (364); she argues that they share similar syntactic and phonetic mechanisms, such as phonetic reduction, syntactic reanalysis, demotivation, fossilization, and conventionalization, but are characterized by different semantic changes. In a detailed study of variant forms of the present perfect in a relic dialect of English (“The Grammaticalization of the Present Perfect in English: Tracks of Change and Continuity in a Linguistic Enclave”), Sali Tagliamonte shows how “the synchronic slice . . . is a striking reflection of an earlier point along the [diachronic] trajectory of grammaticalization of the present perfect” (348). Moreover, she finds that the principles of “layering,” “persistence,” and “specialization” (Hopper 1991) are particularly relevant in the development of the perfect.

The largest number of papers is concerned with the directionality of grammaticalization. According to Minoji Akimoto (“The Grammaticalization of the Verb ‘pray’”), the course of (1) pray [thee, you] to prithee from the fifteenth to the nineteenth century follows many of the recognized principles of grammaticalization, before the ultimate replacement of this politeness form by please. Focusing on one synchronic state of the language, Ursula Lenker (“Soplice and witodlice: Discourse Markers in Old English”) traces the evolution of soplice and witodlice in Old English from manner adverb to truth-intensifying emphaziser to style disjunct and finally to discourse marker on the (meta)textual level, a course of grammaticalization first proposed by Traugott (1995a). Sylvia Adamson (“A Lovely Little Example: Word Order Options and Category Shift in the Premodifying String”) explores the shift in words such as lovely from descriptive adjective (as in cool lovely country) to affective adjective (as in lovely cool country) to intensifier (as in lovely big house) (or from classifying to characterizing to identifying function) and the semantic change from HUMAN PROPENSITY to PHYSICAL PROPERTY to VALUE, as they correlate with position in the noun phrase (leftward movement) and increasing subjectivity (see Traugott 1995b). She notes that some adjectives seem to undergo the opposite direction of change, thus providing an apparent counterexample to the proposed unidirectionality of grammaticalization. Olga Fischer (“Grammaticalisation: Unidirectional, Non-Reversible? The Case of to before the Infinitive in English”) argues that in comparison with Dutch te, which is fully grammaticalized, the English infinitival marker to has undergone a reversal—or degrammaticalization—beginning in late Middle English. She concludes that “the processes underlying grammaticalisation may lead one way as well as another” (152) and that grammaticalization is not a separate and independent process but “can only be discovered with hindsight” (153) (cf. Newmeyer 1998 and Lass, this volume). Extending Fischer’s ideas diachronically, Susan Fitzmaurize (“Remarks on the De-grammaticalisation of Infinitival to in Present-Day American
English”) shows that the occurrence of negative split infinitives in present-day English, as well as the coalescence of to with the quasi-auxiliary in have to and be going to, is further evidence of the degrammaticalization of to.

In “Remarks on (Uni)directionality,” Roger Lass questions a number of the “tentets” of grammaticalization, including the distinction between lexical and grammatical (“much of what we call ‘lexical’ and ‘grammatical’ is based on particular canonical instances in particular canonical languages,” 211), clines of grammatical development (“we could argue then that ‘clines’ or ‘pathways’ have no independent existence or explanatory value; they are post-facto reifications, pure epiphenomena of our notations,” 223), and the “strong” version of unidirectionality (“the claim that all grammatical material is ultimately lexical means that there was a time when all human languages were ‘isolating.’” 216). Even a weak version of unidirectionality, Lass concludes, is “only a set of claims about frequently-observed pathways” (213):

The picture seems rather like this: out of two possible directions, lexical > grammatical, and grammatical > lexical, the first is statistically commoner, so metaphorically ‘preferred.’ (220)

An additional four papers in PCGE bear a somewhat more distant relation to questions of grammaticalization. Concerned more with dating than with process, David Denison (“Combining English Auxiliaries”) provides a wealth of data on the combinatory possibilities of auxiliaries in English as evidence for their grammaticalization from full verbs; he concludes with (quite radically) revised dates for the grammaticalization of the English auxiliaries, ranging from modals, onginnan, and perfect have in Old English; periphrastic do in the fourteenth to fifteenth centuries; passive be in the fourteenth to eighteenth centuries; progressive be in the late eighteenth century; and passive get in the twentieth century. In “Onginnan/Beginnan with to-Infinitive in Ælfric,” Bettelou Los argues that the distinction between the bare infinitive and the to-infinitive following onginnan/ beginnan in Old English relates to auxiliary versus lexical status, bleached (i.e., perfective) versus nonbleached (i.e., ingressive) meaning, and verb-second versus verb-first word order, the former denoting the smooth flow of events and the latter thematic discontinuity. She sees the change from inchoative to perfective meaning as an instance of grammaticalization. Guohua Chen (“The Grammaticalization of Concessive Markers in Early Modern English”) explores in detail the sources of concessive markers, such as (al)though, albeit, for all that, (in) despite (of), notwithstanding, while, and yet. He argues that in Early Modern English, the distinction between factual and hypothetical concessive is not yet fully grammaticalized, though (al)though is preferred for the former and bare if for the latter; what led to the later grammaticalization of the distinction is, Chen admits, “a mystery” (108).
In a detailed and original article, Robert McColl Millar (“Some Suggestions for Explaining the Origin and Development of the Definite Article in English”) argues that the development of the definite article resulted in large part from movement of *that* to a “pure demonstrative,” likely under Norse influence; he pays only brief attention to grammaticalization (277).

Two further papers are rather tenuously connected to the topic of grammaticalization (though the editors try valiantly to relate them to the book’s topic, 11, 31). Rafal Molencki (“Parallelism vs. Asymmetry: The Case of English Counterfactual Conditionals”) surveys the marking of verb forms in the protasis and the apodosis of present and past counterfactual conditionals in the history of English, finding a fluctuation between same and different marking, and Elly van Gelderen (“The Role of Person and Position in Old English”) examines verb movement (to Comp) and loss of agreement markers in Old English, finding loss of such marking more common in the first and second person than in the third person; she finds, conversely, that pro-drop is more common with third person, providing evidence that third person is more highly specified.

The book seems to be carefully edited, though I did find a few small errors, as is inevitable with any such work.

Despite my rather minor reservations about the topical coherence of some of the papers in this volume (though not their ultimate worth), I would recommend this book highly to anyone interested either in the (morphosyntactic and semantic) history of English or in grammaticalization generally, especially in the directionality of change.

**References**


Since the late 1970s, Hans Frede Nielsen has been preparing state-of-the-field surveys of research on the Germanic family of languages and especially on the place of English within that family. Notable among these studies are *The Germanic Languages: Origins and Early Dialectal Interrelations* (1989) and *Old English and the Continental Germanic Languages: A Survey of Morphological and Phonological Interrelations* (1985), both of which were originally written in Danish and subsequently updated and translated into English. Now Professor Nielsen is embarking on an ambitious three-volume history of the English language, surveying publications on the field from the prehistory of Old English down to “the turn of the millennium” (by which he means, presumably, the year 2000 and not 1900), including in his study American as well as British English. (Volume 3 begins in 1776.) The book under review here is volume 1 in this series, the comprehensive title of which will be *A Journey through the History of the English Language in England and America*.

Like Professor Nielsen’s earlier works, *The Continental Backgrounds* provides an inclusive doxastic synopsis distinguished by clarity and care in rendering and integrating the views of the other scholars, but an added element is announced in a section entitled “Fascination Is Our Theme.” That is, in his three-volume history, the author intends to convey, in addition to hard data, “the fascinations of English language history as I perceive them” (5). Bringing as I do a prior commitment to the subject he is treating, I am unable to judge the degree to which he succeeds in this inspirational aspect of his project, but Professor Nielsen certainly writes with verve and with a gift for selecting good examples of any principle he is treating. Routine processing of vast volumes of scholarship on any subject can be a dreary, mechanical exercise, but this is not the case in this book, partly because Professor Nielsen’s own original research is prominent among the scholarly views he is surveying, and so he has a close personal engagement with his subject matter. Indeed, he makes clear that his “three-volume work reflects strong personal views and preferences in the material selected,” but to his great credit, he has the capacity to give fair, faithful summaries of views with which he disagrees. In this respect, his work bears comparison with René Wellek’s (1955-92) multivolume *History of Modern Criticism,*
perhaps the most sustained and successful survey of a field of humanistic research and opinion in our day.

The Continental Backgrounds, like Professor Nielsen’s other volumes, is heavily focused on the internal history of the language, but it includes as well a considerable amount of external history. He devotes several pages, for example, to the questions, “Why did Duke William aspire to become king of England? And how did the Normans get away with conquering the country?” (189). This praiseworthy attempt to cover both kinds of history raises a question about the book’s intended audience. Any educated reader should take an interest in matters of external history, especially when they are presented in the accessible, not to say elementary, form one finds here. But the detailed accounts of Old English morphology and phonology, the concrete discussions of dialectal variation, and the analysis of phonological, morphological, syntactical, and lexical features in comparative entries in the Peterborough Chronicle and the like seem to be addressed to a more specialized audience. It is hard to imagine anyone who has not studied Old English fairly extensively detecting the “fascination” in these topics. Perhaps the students in “intermediate and advanced courses at university level” for whom the author says his book is designed (xi) will find it appropriate to their needs, but my suspicion is that many readers will use the book selectively, the more expert skipping the external history and concentrating on the survey of forms, the less expert enjoying the narrative parts but passing up the detailed analyses of internal linguistic structures.

Cross-references to sections in volume 2 of Professor Nielsen’s Journey through the History of the English Language are frequent and precise, and so work on that volume must be well advanced. The work is written in very fluent English (idiomatic slips like “Full affrication . . . need not to have taken place at once” [107] are rare), and figures, lists, and charts add to the clarity of presentation. Since the author avows that his choice of subjects and scholarship reflects his personal preferences, it would be idle to complain of omissions, but the failure to mention Gneuss’s (1955) Lehnbildungen und Lehnbedeutungen im Altenglischen in the course of the discussion of loan translation can only be an oversight.

References


Just about everyone who takes the Beowulf final comes muttering out of the exam room that he or she could have done better, had the text set for the translation been clean and clear. As in those same seminar discussions when a crux or locus desperatus reared its head, the ambitious first-timer vows to make the text safe for future generations by a cogent analysis of the offending passage(s) in a crisp note or even an extended article. Fortunately, such vows, or “boasts” in the Anglo-Saxon sense, generally diminish directly proportional to the square of the distance from the Schlachthaus. Those who do try the projects of textual criticism—invariably on a local basis—find they tread over old ground, likely repeating old positions formulated in the former Age of Philology, or they finally wilt in the face of the staggering amount of amassed scholarship, or they take up an easy indictment against Klaeber and his benchmark edition, convicting him as a man of his time for his criticism that the poem lacks a steady advance or for his understanding of the characterization of women. Other editors face other charges with similar results: they do not measure up to the felt need or, in the long view, the ever-changing mantra.

Into this fray have stepped Bruce Mitchell and Fred C. Robinson with their edition for the renascent Blackwells of “3182 lines of alliterative verse beginning Hwæt we gardena in geardagum printed often, since Kemble (1833) under the title Beowulf.” It is difficult to imagine two senior scholars who could bring more auctoritas to this text, given their long and distinguished careers in Anglo-Saxon studies. They make it clear in their foreword, however, what their project is: “It is designed to give maximum help to those reading the poem for the first time, and consists of four parts with supporting apparatus—illustrations, facsimile pages, a map, bibliography, and glossary” (vii). They furthermore go on to describe seven aims upon which “we have agreed,” the first and seventh of which are worth noting here:

1. first and foremost, to provide the help needed for an understanding of the text...
of individual passages or of the poem, apart from a short section in which we individually explain our differing views on *Beowulf*. Such problems of judgment and literary interpretation we leave to the student and the teacher. (vii-viii)

The editorial stance would seem to be a model of Cartesian clarity or, less grandly but not less important, a view of the enterprise that meets George Kane’s call of more than two decades ago for “open editions” where editors state their position(s), announce their procedures, and follow through. In one sense, at least, their editorial position on literary interpretation reflects the practical reality that formalism (a.k.a. “new criticism”) has produced: a multiplicity of reasonable, but contrarious if not contradictory, views of the same literary object. Robinson has commented before on the interpreters’ problem in the matter of *Wife’s Lament*. Nevertheless, if the chatnets are any indication, Mitchell and Robinson are being faulted for not having produced the last word and for not replacing Klaeber and above all for not having supported the respective positions of the chatnetting critics passim.

What Mitchell and Robinson have produced is a school text, to use old-style terminology, not a reference text. In doing so, they anticipate a likely classroom setting for user and audience, or teacher and student. Accordingly, for example, the editors create a book where they invite Leslie Webster to provide a significant section on “Archaeology and *Beowulf*.” With its twenty-two illustrations and pointed treatment of the *realia* of the heroic world (helmets, swords, shields, etc.), this section shortens the learning curve for the instructor from a literature background and a beginning student alike, if their take is contextual. Webster reminds all that *Beowulf* is “a poem, not an archaeological textbook,” as she gives an authoritative account of the interdisciplinary nexus, its connections, and its discontinuities ca. the mid-1990s, while the supporting material provides, for example, that classic photo of Trelleborg. The editors themselves present the likely intertextual record. They provide a section on related poems in Old English, deriving texts from volumes 3 and 6 of the *Anglo-Saxon Poetic Records* with translation *en face*, and a section on “documents,” admittedly selective but certainly apposite in providing texts in translation that deal with legendary or mythical materials. Two brief sections treat genealogies and the Geatish-Swedish wars. For those who, having been inspired by the recent reconsideration of the manuscript evidence, might wish to carry this theme forward in their teaching, there are six facsimile pages drawn from the beginning of the poem and from the later work of Scribe B.

All this background information has its mate in Part One, the Introduction. With their intended audience in mind, Mitchell and Robinson give a synoptic treatment of the major points from manuscript features through date and place of composition, language, structure, the literary topics of style and subject, and, finally, their “two views” on the poem. Partisan commentators may not finally agree with this re-
viewer’s judgment, but by and large Mitchell and Robinson are even-handed in their approach to some knotty problems. They seem able to make a point for beginners (viz., Old English poets did not provide their poems with titles) and to move on to the more abstruse, such as the problem of fitt numbering, without changing registers. Along the way, the first-timer begins to learn that the composition of Beowulf has no contemporary analog to creative writing, as in the phrase “date of composition,” which gets defined as “the approximate date when the tradition-bound, accretive process we call Beowulf assumed more or less the shape it has in the copy we have in Cotton Vitellius A. xv.” The language section contains the sort of humane philology that specialists can only practice, if they wish to reach contemporary, language-poor audiences. The teaching voice is very much in evidence in the more literary discussions, where the commentary slowly begins to emerge as a dialogue between the editors, ultimately yielding their separate, solo takes on the poem. In fact, the two views are really a duet, carrying strains of Tolkien and apparent differences in emphases, not substance, while confirming that scholars with great technical expertise can produce broadly humane criticism, moving in its appeal and relevant to our time. All in all, the Introduction can prove to be an excellent refresher for the sometime Beowulf teacher who comes back to the poem only so often and who needs to become current in a hurry. One might want to have similar discussions of women, monsters, and even the razzle-dazzle of computer approaches, all of which have spawned countless papers at conferences in recent years to the extent that those 3,182 lines known as Beowulf might in some quarters be considered to be really about Grendel, his dam, and a most reluctant dragon. But the surge in the subject in these areas has been very recent, coming forward at about the time this edition was making its way through the press.

In the spirit of full and open disclosure, Mitchell and Robinson cite their principles and procedures of their edition briefly as a headnote to the text proper and then as a separate part “How We Arrived at the Text.” Two statements stand out in high relief: “Our editorial approach can be described as conservative in that we prefer to retain the manuscript reading where possible” (165). Then, after a primer on the various kinds of textual problems, Mitchell and Robinson address this metaphor to the reader: “Reading through these sections should have given you some idea of the hidden currents, whirlpools, reefs, and shipwrecked theories, which are to be found beneath the surface of the river of poetry we present to you” (171). This implicit introduction to current editorial theory, which features noninterventionism and indeterminacy, then moves on to a specific consideration of punctuation. With reservations expressed, Mitchell and Robinson offer modern punctuation in their text, seeking a minimal system so as not to impede the flow of the verse paragraph and so as not to challenge the first-time user unduly. The exclamation mark is a case in point. Klaeber, characterized as “an hysterical punctuator,” uses the exclamation mark some fifty-five times, but Mitchell and Robinson retain only four of these and
add three of their own. There are, of course, many other issues in punctuation, which impinge upon meaning and its subtleties, and these are sketched as well. It is well known that Mitchell has cogent views on “the dangers of disguise” in Old English editions that feature modern punctuation. Mitchell and Susan Irvine have now issued *Beowulf Repunctuated, Old English Newsletter Subsidia 29* (2000).

The layout and design of the text proper have their helpful features: set in 10.5 on 12-point Ehrhardt, the typeface is easy to read, while margins seem spacious enough for brief comments in the approximately 6 × 9 trim size (xeroxable, of course, to a larger size for those who want more margins for notes). The editors provide bracketed titles to each fitt, and notes are below the text. The first deck of notes is textual and laconic, given where the editors depart from the manuscript (thus line 499 of the text gives ‘Unferð’ and the note ‘HVN ferð’), and without the sort of interpretation history Klaeber gives. The second deck of notes gives translations, interpretations, and discussions of particular points without lengthy bibliographical pursuits. The interpretation of specific points will, as is to be expected, inspire the most controversy. The explanation of line 2820, ‘soolfæstra dom,’ is correct at the root, and the editors give two alternate translations, but some will want to know more. Likewise line 1931b, ‘mod þryðowæg,’ about which much has been written, receives the ingenious explanation where ‘wæg’ takes on the meaning ‘weighed’ and its true subject is Hygd. The notorious ‘ealuscerwen’ receives comparatively full treatment because it is referred back to the section treating problems in editing the text, where it is an exemplary and illustrative difficulty. The editors supply long marks on vowels, apparently without signaling this quiet move, but they do call attention to the vagaries of word spacing in the manuscript.

Spot checks in the text proper, made against Zupitza-Davis in EETS OS 245, reveal line 419b, ‘searum,’ should really be ‘searwum’ (likewise in the glossary). The editors navigate the shoals of the last damaged page quite well—only 3171b poses a minor problem. The editors regularize to ‘cyning,’ noting that the manuscript reads ‘scyning.’ Zupitza-Davis reads ‘kyning’ in the manuscript as does Klaeber. Through the mists of EETS OS 245, one can make out a letter form much like a tall s before c, but Scribe A at lines 619 and 665 uses an exuberant k-form, which admittedly lacks the curl to the right at the top of the descender. Despair may very well be in the details for “the late hand” (presumably Scribe B), according to Zupitza-Davis, is responsible for the travesty ‘lachaman’ where ‘lichaman’ is clearly meant. The pursuit of such fine points is an expert’s province, as Browning unkindly noted in his poem about a certain grammarian.

Mitchell and Robinson have made their objectives clear at the outset and have, by and large, attained them. The deep, strategic goal is to convert more readers to Old English, and to do so they have built a bridge to that same public that greeted Seamus Heaney’s new translation and now Roy Liuzza’s, which takes reading from Mitchell and Robinson. Historically, the field of Old English has never had the
same strong pedagogical foundation as the Classics, and the standard move from beginning Old English to *Beowulf* is too formidable for some students and too impossible for some curricula, especially in the United States. Mitchell and Robinson have made it possible for many to experience what Borges did, as they cite him on page ix: to study “the blunt-tongued Anglo-Saxons,” to contemplate the self, and to see that “The universe waits, inexhaustible, inviting” (as Alastair Reed translates).
BOOKS RECEIVED


INDEX

to

JOURNAL OF ENGLISH LINGUISTICS

Volume 29
Number 1 (March 2001) pp. 1-96
Number 2 (June 2001) pp. 97-196
Number 3 (September 2001) pp. 197-288
Number 4 (December 2001) pp. 289-388

Authors:

BOTLEY, SIMON, and TONY McENERY, “Proximal and Distal Demonstratives: A Corpus-Based Study,” 214.
BUTTERS, RONALD R., “Chance as Cause of Language Variation and Change,” 201.
CAPPELLE, BERT, “Is out of Always a Preposition?” 315.
GINER, MARÍA F. GARCÍA-BERMEJO, and MICHAEL MONTGOMERY, “Yorkshire English Two Hundred Years Ago,” 346.
KIM, SUKSAN, “Lexical Phonology and the Fricative Voicing Rule,” 149.
KIRK, JOHN M., see Montgomery, M. B.
KRETZSCHMAR, WILLIAM A., Jr., “Using Computers in Linguistics: A Prac-
LEE, DAVID Y. W., “Defining Core Vocabulary and Tracking Its Distribution
across Spoken and Written Genres: Evidence of a Gradience of Variation from
the British National Corpus,” 250.
LEISTYNA, PEPI, “Discourse in Late Modernity: Rethinking Critical Analysis,
by Lilie Chouliaraki and Norman Fairclough” [Book Review], 183.
MARKUS, MANFRED, “Linguistic Commercialism in and around the Paston and
Cely Letters: An OED and Corpus-Based Approach,” 162.
McENERY, TONY, see Botley, S.
MINKOVA, DONKA, “The Cambridge History of the English Language, Vol. 3,
edited by Roger Lass” [Book Review], 83.
MONTGOMERY, MICHAEL, see Giner, M.F.G.
MONTGOMERY, MICHAEL B., and JOHN M. KIRK, “‘My Mother, Whenever
She Passed Away, She Had Pneumonia’: The History and Functions of when-
ever,” 234.
MOSES, STEPHANIE, “Language Teaching Awareness, by Jerry Gebhard and
Robert Oprandy” [Book Review], 285.
MUKHERJEE, JOYBRATO, “Principles of Pattern Selection: A Corpus-Based
Case Study,” 295.
OH, SUN-YOUNG, “A Focus-Based Study of English Demonstrative Reference:
With Special Reference to the Genre of Written Advertisements,” 124.
PARKER, FRANK, “Syntactic Structures Revisited: Contemporary Lectures on
Classic Transformational Theory, by Howard Lasnik” [Book Review], 179.
RENOUF, ANTOINETTE, see Bauer, L.
RILEY, KATHRYN, “Routledge Dictionary of Language and Linguistics, by
Hadamod Bussman; translated and edited by Gregory P. Trauth and Kerstin
Kazzai” [Book Review], 282.
ROBINSON, FRED C., “The Continental Backgrounds of English and Its Insular
Development until 1154, by Hans Frede Nielsen” [Book Review], 386.
Bryan A. Garner” [Book Review], 279.
SWALES, JOHN M., “Metatalk in American Academic Talk: The Cases of point
and thing,” 34.
by Bruce Mitchell and Fred C. Robinson” [Book Review], 388.

Articles:

“Chance as Cause of Language Variation and Change,” Butters, 201.

“Demonstratives in English: A Corpus-Based Study,” Botley and McEnery, 7.


“A Focus-Based Study of English Demonstrative Reference: With Special Reference to the Genre of Written Advertisements,” Oh, 124.


“Is out of Always a Preposition?” Cappelle, 315.


“Lexical Phonology and the Fricative Voicing Rule,” Kim, 149.

“Linguistic Commercialism in and around the Paston and Cely Letters: An OED and Corpus-Based Approach,” Markus, 162.

“Metatalk in American Academic Talk: The Cases of point and thing,” Swales, 34.

“My Mother, Whenever She Passed Away, She Had Pneumonia’: The History and Functions of whenever,” Montgomery and Kirk, 234.

“Principles of Pattern Selection: A Corpus-Based Case Study,” Mukherjee, 295.

“Proximal and Distal Demonstratives: A Corpus-Based Study,” Botley and McEnery, 214.


“Yorkshire English Two Hundred Years Ago,” Giner and Montgomery, 346.

**Book Reviews:**


“Language Form and Language Function, by Frederick J. Newmeyer,” Cureton, 72.


“Pathways of Change: Grammaticalization in English, edited by Olga Fischer, Anette Rosenbach, and Dieter Stein,” Brinton, 382.

“Qualitative Methods in Sociolinguistics, by Barbara Johnstone,” Johnson, 374.
“Syntactic Structures Revisited: Contemporary Lectures on Classic Transformational Theory, by Howard Lasnik,” Parker, 179.