

THE DATA OF HISTORICAL LINGUISTICS: SOURCES FOR THE  
RECONSTRUCTION OF PRONUNCIATION FROM WRITTEN RECORDS\*

Charles V.J. Russ

The historical linguist is somewhat at a disadvantage when compared with his synchronic counterpart as regards the question of data at his disposal. Firstly, the historical linguist has had his data selected for him by circumstances beyond his control, and the amount of data is much smaller than that available to the synchronic linguist. Secondly, the data usually only represents the language of one particular social, political or religious class in society, mostly the ruling class. Thirdly, the data comprises exclusively written sources (there is naturally no recourse to native speakers). Given these limitations, it is surprising, and gratifying, to see how much these written sources have yielded up in the last hundred years. Theories of historical linguistics have been proposed, changed and refined in many ways, but I do not intend to deal with individual theories in this paper. My main purpose is rather to examine the principles which may be used in the interpretation of written sources which I consider to be the basic data of historical linguistics. Correspondences such as Old High German (OHG) *th*, New High German (NHG) *d* form the skeleton on which the various theories of historical linguistics have been built. (I shall assume that such correspondences represent a meaningful relationship in time.) It has been usual to assume that a reconstructed value of some kind can be given to an initial symbol, e.g. OHG *th*, whereas the value of the terminal symbol, e.g. NHG *d*, can be ascertained directly, although in some cases there *is* controversy about the terminal value of symbols, cf. the discussion of vowel quality and quantity in English and German.

If there is to be such a discipline as historical linguistics then its principles, if they are to be of any interest and value, should be universally applicable. However the principles for the reconstruction of pronunciation from written records are rarely made explicit and open to scrutiny and evaluation. Recent books on historical linguistics barely mention them: Samuels (1972:4) and Anderson (1973:16) deal with them in one sentence and Anttila (1972:36) deals only with inverse spellings. Older scholars such as Jespersen and Joseph Wright, however, made more explicit statements about what they called 'the chief sources for ascertaining the approximate pronunciation of the speech-sounds of our ancestors', (Wright 1925:3). The types of evidence Jespersen and Wright suggest, which coincide almost exactly, comprise statements by phoneticians and grammarians (including foreigners teaching English to their fellow countrymen), rhymes, and occasional phonetic spellings. In addition to these Jespersen mentions puns. These sources were taken over by later scholars; Bloomfield mentions them all but only very briefly (1935: 294-6). Generative phonologists too have dealt with reconstruction but have not been explicit about their methods. Chomsky and Halle, for instance, assume the traditional phonetic analysis of the Middle English vowel system and only use the works of phoneticians through the centuries to reconstruct changes (1968:249-89). There have of course been many discussions of individual reconstructions in particular languages, but on the whole the discussion of methods of reconstruction has been sparse. A notable exception to this has been Herbert Penzl, whose main concern has been with the history

of standard German. He discusses the following sources for the reconstruction of pronunciation: evidence from spelling, i.e. occasional or inverse spellings, evidence from loan words, comparative evidence, which, in his opinion, includes not only evidence from other languages of the same family but also from the dialects of the same language (Penzl 1957:197-200; 1972:72-9). In later works he adds another source, the comparison of the assumed pronunciation with previous, or more importantly, subsequent sound changes: 'By comparing the assumed sound changes and sound values in subsequent developments we can not only attempt to determine phonemes but allophones including their concrete phonetic characteristics' (1971:21, author's translation). Professor Sidney Allen in his detailed reconstructions of the pronunciation of Latin and Greek (1965, 1968) assumes all the sources we have mentioned already, as does Gimson (1970:73-9) in his *Introduction to the Pronunciation of English*.

Many scholars have therefore assumed that it is quite feasible to reconstruct pronunciation, indeed Allen is quite positive about it: 'The degree of accuracy with which we can reconstruct the ancient pronunciation varies from sound to sound but for the most part can be determined within quite narrow limits' (1965:vi). Lass and Anderson in their recent book *Old English Phonology* are more sceptical: 'As far as we can tell, historical investigation does not permit us to recover much more than grossly binary (i.e. classificatory) specifications..... We must attempt, as we have all along here, to achieve reasonably full specification at the phonological and systematic phonetic levels; but when it comes to pronunciation..... we must simply say that we have no relevant information, and cannot conceive how we might get it, short of a time machine,' (1975:203). These remarks are primarily concerned with the reconstruction of vowel quantity and quality in Old English. Elsewhere in the book they *do* try and specify in greater detail phonetic values for the letters *l*, *r* and the digraph *ie* (83-9; 122-9). Wrenn in his article 'The value of spelling as evidence' (*TPS*, 1943), although fundamentally agreeing that pronunciation can be reconstructed, constrains every principle with caveats. McIntosh (1956) seems even more sceptical and wants to emphasize that there are orthographic variations which do *not* reflect any variation, e.g. *þ* and *th* in ME. His remarks provide a healthy antidote to anyone who might be tempted to always identify any orthographic change with a sound change, but the latter view was repudiated as long ago as 1892 by Kauffmann: 'We have furthermore come to realize....that sound changes are in no way completely identical with orthographic changes' (1892:243, author's translation). In this paper I shall confine my remarks to the reconstruction of segmental sounds, although it is quite possible to reconstruct prosodic, or supra-segmental, features, as shown by Allen's study *Accent and Rhythm, Prosodic Features of Latin and Greek: A Study in Theory and Reconstruction* (1973).

It might be argued by some that a discussion of methods, or principles, is a mere discussion of 'discovery procedures' and as such is pre-theoretical and outside the scope of any theory of linguistics. It is my contention, however, that this is not so, but that the methods of reconstruction fundamentally affect the statements made about the initial values of segments which then act as the input for sound change rules or correspondences. Lehmann, (1973:146) for instance, assumes -- he admits he cannot substantiate his claim -- that ME /t/ was dental [t̪] and has subsequently become alveolar in modern English. (He assumes that /d/ and /n/ were also dental.) The traditional view assumes that ME /t/ has not undergone any change and thus

was alveolar in ME. If we assume with Lehmann that ME /t/ was dental then English has undergone a change among the voiceless stops of dental to alveolar; his description of the history of English will contain an additional sound change to the traditional descriptions which assume that ME /t/ has not changed. This is surely an important difference and is due to the assumption of different initial values for ME /t/. A much more important example which shows how reconstruction is crucial in defining the initial stages of sound changes is the one discussed by Lass in: 'What kind of vowel was ME /a/?' (1973). He explicitly states that 'the entire history of the West Germanic low vowel \*/a/.....in English will vary according to our reconstruction of the shape of the vowel systems at any point' (1973:61). If /a/ is assumed to be front then only raising is required to account for the modern English value, if however /a/ is not front but back, then a sound change rule involving both fronting and raising is needed to convert [ɑ] to [æ]. The choice of the initial phonetic value of a symbol has important consequences for the description of the sound changes in the historical development of languages. 'Surely principled answers to questions like whether or not a sound change has occurred, and if it has where, are vital to the construction of a linguistic history' (1973:62). To my mind these two examples show clearly the importance that the choice of an initial phonetic value can have in describing sound changes which have taken place in a language.

In the main part of this paper I should like to scrutinize the principles which have been proposed to help reconstruct pronunciation and examine their strengths and weaknesses with a view to putting them on a hierarchical scale of usefulness.

### 1. *Statements on pronunciation by contemporary phoneticians and grammarians.*

This is the nearest one can come to direct evidence for pronunciation in the past, but even so it is subject to many limitations. Evidence like this is not available for all languages at all times. The classical languages, particularly Sanskrit, are well endowed in this respect, but evidence for the other European languages does not really start till after the Renaissance, when, with the decline of Latin, the vernacular languages began to be widely used in writing. There is, however, one isolated example in the Germanic languages of a contemporary phonetic description of a language at an earlier date. This is the account of the spelling and pronunciation of twelfth century Icelandic by the so-called First Grammarian (edition by Benediktsson 1972). His phonetic descriptions are for the most part quite easily interpreted. For instance he noticed that the letter *n* before *g* was pronounced as a velar nasal: 'The *n* which comes before a *g* is spoken less in the nose and more in the throat than the other *n*'s, because it receives some slight admixture from the *g*', (Benediktsson 1972:237). Although there is some useful material, in the sense of being clearly interpretable, in the works of many phoneticians and grammarians, the majority of comments are unclear or ambiguous. Lass cites John Hart's description of *a* as being pronounced 'with wyde opening the mouth, as when a man yauneth' as being ambiguous and not showing whether the sound was a front or back vowel (1973:64). German phoneticians of the sixteenth century used the term *hart* 'hard' in such statements as: 'The *b* and *p* are also the same, except that the *p* is harder than the *b*, thus *t* is also harder than *d*' (Müller 1882:130). There was clearly a distinction between *p* and *b* but it is not clear from the statement whether it involved voiceless vs. voiced, fortis vs. lenis, or aspirated vs. unaspirated. Sometimes

grammarians' descriptions are simply wrong: the front rounded vowels in German are often described as 'diphthongs' (Müller 1882:66), or consonants written double are considered to be pronounced long (Müller 1882:100). Sources such as these have sometimes provided useful information but they must be used with caution remembering Jespersen's warning: 'It would be an extremely grave error to suppose that every little notice found in an old grammar about the pronunciation of such and such a word is the exact truth....' (1909, 1:9).

## 2. Loan Words.

If it seems that the pronunciation of the loan words has not changed since the time they were borrowed, then a reconstruction of a more concrete phonetic nature may be assumed. The initial affricates of the French loan words *chant*, *chain*, *judge*, *joy* in modern English are probably similar to the pronunciation of initial *ch* and *j* in Old French. In modern French the affricates [tʃ] and [dʒ] have been simplified to the fricatives [ʃ] and [ʒ]. OHG *s* has been reconstructed as a [ʃ]-like sound (Braune/Mitzka 1961: para. 168) from the fact that OHG loans in Slavonic languages have been rendered by [ʃ] and not by [s], German *Stollen* 'gallery' (in a mine), Czech *štola*. The Latin loan word *keisur* in OHG, with *ei* pronounced as a diphthong, cf. NHG *Kaiser*, is often used to show that when it was borrowed by OHG, Latin *ae* was still a diphthong (Allen 1965:60). It may have been a diphthong, but what was its exact value? The use of loan words, except when we can reasonably assume that their pronunciation has changed either little or not at all, is fraught with difficulties. They must only be used with great caution since in being borrowed, the sounds may have changed their articulation to become better integrated into the phonological pattern of the borrowing language, and also, as for instance in the case of Latin *caesar*, OHG *keisur*, we are working with two unknowns and not just one.

## 3. Phonological development in the language itself.

In some cases, sounds which contrast in a present-day language do not contrast orthographically in an older period of the language. Unless there is good reason for us to assume that they have resulted from the split of one sound then it can be assumed that the phonemic contrast in the modern language also existed at earlier stages. A good example of this is the reconstruction of short and long vowels in MHG (in most medieval stages of European languages long vowels are not marked especially, (Allen 1965:64)). In writing no distinction was made between the vowel in MHG *wibe*, *libe* (dat.sg.), *siben*, *riben* (the handbooks, dictionaries and normalized texts write the long vowels with <sup>^</sup>). However, if we examine their reflexes in NHG, we find that *wibe* 'woman', *libe* 'life' are represented by diphthongs, *Weibe*, *Leibe*, whereas *siben* 'seven', *riben* 'rubbed', are represented by long vowels [zi:bən], [ri:bən]. There is no evidence to show us that the phonetic values in NHG resulted from a split in the one original MHG sound. Since the reflexes are distinct in NHG we can also assume them to have been distinct in MHG. On the other hand the modern English contrast /u/ : /ʌ/, as in *put* : *putt*, is not assumed to have existed in ME. The vowels of ME *ful*, *putte*, *up*, *thus*, *hunte*, were all written with the same sign. In modern English, however, *full*, *put*, are pronounced with [u], but *up*, *thus*, *hunt* with [ʌ]. According to the principle we have set up, we can assume the /u/ : /ʌ/ contrast for ME if

there is no evidence of a phonemic split of ME /u/. In this case, unlike the MHG case just mentioned, there is however evidence of a phonemic split. The distribution of the phoneme /u/ is restricted in modern English. Words in modern English written with *u* and pronounced [u] usually begin with a labial, e.g. *pull, full, put, butcher* and often end in *l*, whereas [ʌ] does not appear so often after initial labials (Gimson 1970:107ff.). It has been assumed that ME /u/ developed a lowered, less rounded allophone [ʌ] except after labials. At one time [u] and [ʌ] were in complementary distribution. Due to the shortening of ME *ō*, which had by then been raised to *ū* before *k*, as in *book, look, cook*, minimal pairs such as *look : luoek* /u/ ≠ /ʌ/, were created (Kurath 1964:96). This method of reconstruction is quite effective for projecting contrasts back into the past, even if there is little or no orthographic evidence for the distinction in ME or MHG. In this way phonemic oppositions can be established, but not the exact phonetic realization of the phonemes.

#### 4. Spelling conventions and variants.

In many texts, particularly in Early NHG, the spelling seems to be in a great state of flux. Many words are spelt differently even in the same sentence. The same sign is used randomly, varying with different ones in different words. For example, the NHG word *Zeit* 'time' is written in one text: *zeit, zeyt, zit, zyt* and the NHG word *Bein* 'leg' is written: *bein, beyn, bain* (Philipp 1968: 4f. and 93ff.). There is an overlapping of the signs *ei, ey* in the spelling of these two words. In some words *ei, ey* alternate with *i, y* and in other words with *ai*. If, instead of looking at the individual signs, we look at groups of signs then the fluctuation seems less random and more capable of regular description. In the example just quoted, we have what might now be called two graphemic 'variables', one comprising the set //i, y, ei, ey// and another comprising the set //ei, ey, ai//. Since there are two sets of signs, even though *ei* and *ey* occur in both of them, it is assumed that we are dealing with two different sounds which are in opposition. This is further supported by the fact that the reflexes of MHG *î*, represented by the set //i, y, ei, ey//, and the reflexes of MHG *ei*, represented by the set //ei, ey, ai//, do not rhyme.

Changes in spelling most often occur when phonemic mergers take place. In the case of the merger of MHG *î* and *ei*, the sign *ei* becomes predominant. The first sign of a phonetic merger is that two signs which hitherto have been kept carefully apart become used wrongly, the ongoing merger of MHG *s* and *z* can be seen when *das* is written for *daz*, and *allez* for *alles* (Schulze 1967:38f.). This type of evidence is known as inverse spellings (Jespersen 1909, 1, 4). At the beginning of the orthographic recognition of a merger this may not happen very frequently and then such 'slips' are often called occasional spellings. Such spelling 'mistakes' usually show any phonemic change which involves conditioned merger, and changes in the incidence of phonemes (Penzl 1957:200f.). Other phonemic changes such as shifts and splits may also be reflected in changes in orthography, but not to the same extent as mergers are. Spelling conventions and variants provide a mass of data for the reconstruction of sound changes and pronunciation, but again chiefly of which sounds contrast and which do not.

### 5. *The evidence from rhymes and puns.*

The evidence from rhymes and puns tells us which sounds were pronounced in the same way or considered to be pronounced in the same way, and by implication those sounds which were not rhymed were not pronounced the same. Like the other methods of reconstruction which we have considered rhymes tell us little about the actual phonetic realizations of the sounds but only whether they contrast with other sounds or not. End rhymes can usually tell us about stressed vowels, intervocalic consonants and final consonants and vowels. Rhyming practice has been one of the main supports for the view that MHG possessed two short *e* sounds, a close [e], *besser* 'better', historically from Germanic short *a* by *i* mutation, and an open [ɛ], *essen* 'to eat', representing Germanic short *e* (Zwierzina 1900:250-316). Most MHG poets did not rhyme these two sounds, and even though they are both spelt *e* it is assumed that MHG had two phonemes /e/ and /ɛ/. Caution must be displayed in using rhymes to ascertain which sounds were pronounced the same since some rhymes are only rhymes to the eye and not to the ear, e.g. in English, *love:move*, *pant:want*. This practice seems to have developed in the sixteenth century and thereafter (Wrenn 1943:36). Puns perform a similar task to rhymes; they indicated which words are considered to have the same pronunciation. The use of puns for the reconstruction of sound changes and pronunciation has chiefly been applied to English. K&Okeritz (1953:53-157) cites many puns from Shakespeare.

### 6. *The evidence from related languages and their dialects.*

The methods or principles discussed hitherto have chiefly been able to tell us whether certain orthographic contrasts also reflected phonological surface or phonemic contrasts, only the use of the comments by grammarians and in some cases the use of loan words could give us greater, though limited, phonetic detail. In practice most of the positive phonetic reconstruction of pronunciation has come from other languages in the same language family or from the modern dialects, or even the standard language of the language concerned. The digraph *th* in OHG is considered to represent a voiceless interdental fricative [θ] since this pronunciation has been retained in modern English and Icelandic. Again the evidence may sometimes be controversial: what value did Proto-Germanic initial *g* have? English has a palatal glide, Dutch a voiced or voiceless velar fricative, and other languages a stop or fricative whose point of articulation varies according to the following vowel. Was it a fricative or a stop, palatal or velar?

Penzl (1957:199) calls the evidence from modern dialects 'neocomparative' in contradistinction to comparative evidence from languages in the same language family. The MHG example cited in the section on rhymes showed that in MHG /e/ was different from /ɛ/, and if we examine modern Upper German dialects we can see that this contrast is still maintained as that between a half-close [e] and a half-open [ɛ] or [ə]. Thus the modern dialects can give us phonetic detail to fill in the phonological contrast we have established on the basis of rhymes. On the other hand the lack of contrast we assumed in ME for the modern English contrast /u/ : /ʌ/ on orthographic grounds is supported by the North Midland and Northern dialects in England which do not have this contrast (SED II.1.2 'stubble'). Recently Lass has investigated ME short /a/, trying to ascertain whether it was front or back. This question can be resolved, he says, 'by means of available but neglected

evidence. The one source of information that might really help, and by far the most important in a reconstructive venture like this, is the one that has not been systematically employed: the actual range of reflexes of ME /a/ in a wide range of modern dialects' (1973:67). He comes to his conclusion, he calls it 'a substantive claim', that '...ME *a* ... must be reconstructed as an open front vowel' (1973:81). Appeal has always been made to the phonetic values of sounds in the dialects in the study of German (Kauffmann 1917:45), and with the data from the SED available this must now be taken into account in the history of English. However, one caveat must be introduced; one cannot simply project modern phonetic values in dialects on to symbols in medieval manuscripts. The dialects may have changed, particularly in the case of German -- this is true of those dialects which are in close contact with other languages, e.g. in the south of Switzerland. Moulton has shown (1941: 19f.) that the retention of final vowels in the Swiss Valais, which is usually taken as simply being a very archaic trait, is in fact due to the weak stress which comes from the Gallo-Romance substrate. Given this caveat, modern dialects do provide probably the best laboratory for the historical linguist who wants to reconstruct pronunciation.

#### 7. *The naturalness of reconstructions and sound changes.*

A final principle, one which is not explicitly stated by the traditional scholars we have cited, but implicit in their work, is that of the naturalness of sound changes. When reconstructing the initial value (I) for an orthographic symbol, knowing its terminal value (T), the reconstructed value must be such that the change from (I) to (T) is a 'natural' change. Thus [i] is not likely to change to [p]. The use of this principle can be seen in Bailey's reconstruction of *zeta* in Ancient Greek, where he challenges scholars to prove that changes like *zd* > *z(z)*, *dz* > *dd* are possible (Bailey 1968:177f.). Lass and Anderson (1975) also work with this principle: 'It is a commonplace in both synchronic and diachronic linguistics that certain kinds of rules or sound changes are 'natural', and others less so. And in practice this conviction has sufficient force for us often to opt for a particular underlying form or reconstruction on the ground that -- given the observable output, and a choice of say two processes for getting it -- one alternative seems 'highly valued', and the other less so.' (1975:148). What exactly *is* a natural sound change has not yet been spelt out completely, but this does not affect the general point here that any 'naturalness' that is presupposed will constrain reconstructions.<sup>1</sup>

Viewed on a scale of usefulness, the use of rhymes, puns and orthographic variants allows us to establish a class of contrasting phonological units, while use of modern dialects and other languages enable us to put more precise phonetic values to these units. We may be helped in this latter to a lesser extent by statements of grammarians and the use of loan words. Finally every reconstruction will have to obey the constraint that it is 'natural', i.e. the assumed change it undergoes belongs to the class of possible sound changes.

I have tried to show that since the data of historical linguistics consist largely of written sources, the principles for the analysis and reconstruction of pronunciation as initial values for sound changes are important and worthy of discussion within a general framework rather than

in individual cases of reconstruction. To my mind this is a worthwhile and challenging task for the historical linguist.

## NOTES

- \* My thanks are due to my colleague Dr John Green who read through this paper before it was delivered and made some useful suggestions on points of detail which I have followed. For the contents of the paper I myself must bear the full responsibility.
1. It was pointed out in the discussion by Roger Lass that 'natural' should be taken to mean 'natural' in a particular language, e.g. 'natural' in English.

## REFERENCES

- Allen, W.S. (1965) *Vox Latina*. Cambridge, C.U.P.  
 (1968) *Vox Graeca*. Cambridge, C.U.P.  
 (1973) *Accent and Rhythm, Prosodic Features of Latin and Greek: A study in theory and reconstruction*. Cambridge, C.U.P.
- Anderson, J.M. (1973) *Structural Aspects of Language Change*. London, Longman.
- Anttila, R. (1972) *An Introduction to Historical and Comparative Linguistics*. London-New York, Macmillan.
- Bailey, C.-J. (1968) 'The Pronunciation of zeta in Ancient Greek', Chicago Linguistics Society, *Papers from the fourth regional meeting*:177-93.
- Benediktsson, H. (ed. 1972) *The First Grammatical Treatise*. Reykjavik, University of Iceland Publications in Linguistics.
- Bloomfield, L. (1935) *Language*. London, Allen and Unwin.
- Chomsky, N. and Halle, M. (1968) *The Sound Pattern of English*. New York, Harper and Row.
- Gimson, A.C. (1970) *An Introduction to the Pronunciation of English*. London, Arnold.
- Jespersen, O. (1909) *A Modern English Grammar on Historical Principles; Part 1: Sounds and Spellings*. London, Allen and Unwin.
- Kauffmann, F. (1892) 'Über althochdeutsche orthographie', *Germania*, 37:243-65.  
 (1917) *Deutsche Grammatik*, Marburg, Elwert.
- Kökeritz, H. (1953) *Shakespeare's Pronunciation*. New Haven, Yale University Press.
- Kurath, H. (1964) *A Phonology and Prosody of Modern English*. Heidelberg, Winter.
- Lass, R. (1973) 'What Kind of Vowel was Middle English /a/, and What Really Happened to it?' Edinburgh University Linguistics Department, *Work in Progress* 6:60-84.
- Lass, R. and Anderson, J.M. (1975) *Old English Phonology*. Cambridge, C.U.P.
- Lehmann, W.P. (1973) *Historical Linguistics*. New York, Holt.

- McIntosh, A. (1956) 'The Analysis of Written Middle English', *TPS*:26-55.
- Moulton, W.G. (1941) *Swiss German Dialect and Romance Patois*. Baltimore, LSA, *Language Dissertation* 34.
- Müller, J. (1882) *Quellenschriften und Geschichte des deutschsprachlichen Unterrichts*. Repr. 1969 with an introduction by M.Rössing-Hager, Wissenschaftliche Buchgesellschaft, Darmstadt.
- Penzl, H. (1957) 'The Evidence for Phonemic Changes', *Studies Presented to Joshua Whatmough*. The Hague, pp.10-24.
- (1971) *Lautsystem und Lautwandel in den ahd. Dialekten*. Munich, Hueber.
- (1972) *Methoden der germanistischen Linguistik*. Tübingen, Niemeyer.
- Philipp, M. (1968) *Phonologie des graphies et des rimes: l'alsacien de Thomas Murner*. Paris.
- Samuels, M.L. (1972) *Linguistic Evolution*. Cambridge, C.U.P.
- SED (=H.Orton ed. 1962ff) *Survey of English Dialects. Basic Material*. Leeds, Arnold.
- Schulze, U. (1967) *Studien zur Orthographie und Lautung der Dentalspiranten S und Z im späten 13. und frühen 14. Jahrhundert*. Tübingen, Niemeyer.
- Wrenn, C.L. (1943) 'The Value of Spelling as Evidence', *TPS*:14-39.
- Wright, J. & E.M. (1925) *An Elementary Historical New English Grammar*. Oxford, O.U.P.
- Zwierina, K. (1900) 'Mhd. Studien 8: Die e-Laute in den Reimen der mhd. Dichter', *Zeitschrift für Deutsches Altertum*, 44:248-316.