SOME NOTES ON THE ORIGIN OF MIDDLE ENGLISH /a/*

ARLETA ADAMS-KASZCIK

Adam Mickiewicz University, Poznań

The development of West Germanic a¹ in English presents diachronic linguists with more problems than the history of any other short vowel. This is because the vowel has been exceptionally unstable: not only has it changed more frequently than other vowels, but also, if we look at the changes in their totality, their effects appear strikingly circular. This means, among other things, that MnE reflexes of WG a are not radically different from their source, unlike, for instance, MnE reflexes of ME long vowels.

The history of WG a in the RP and related varieties of southern British English can be sketched as follows:

The first split represents the Anglo-Frisian Brightening, the second shows the effects of breaking and velar umlaut; the first merger is the Late Old English monophthongization of æa, the second — the Early Middle English coalescence of æ and a. Next comes the Early New English raising of ME a to æ, then the lengthening of the latter before /f, ð, s, ns/ in the late 17th century and, finally,

* The paper was presented at the XVI Annual Meeting of the Societas Linguistica Europaea, held in Poznań from 19-21 August 1983.

¹ Five types of notation are used throughout, with the following reference:
- an etymological class or the class of reflexes of a change
- a phoneme
- an allophone or the IPA value (if phonemic status is irrelevant)
- a reconstructed sound
- a graph.

The peculiarity of these developments rests in the recurrence of one particular segment, namely ae, which has somehow been the "favoured" reflex throughout the history of the dialects in question. In other dialects, however, the "centres of gravity" — to use Lass' term — appear to have been different; thus, e.g., a has on the whole dominated in Eastern Border Scots type (Lass 1977:18).

As we can see, the diachronic instability of WG a has been accompanied by great diatopic differentiation. This is particularly conspicuous in present-day varieties of English; in fact, attempts at describing this synchronic variation in sociolinguistic terms have not been infrequent (see, e.g., Labov 1966, Trudgill 1974 and Romaine 1982a).

Of all the issues touched upon so far, our main concern will be, in general, the reasons for the quite exceptional instability of the reflexes of WG a (in time rather than space) and, in particular, the motivation for the merger which took place some time in Early Middle English.

In trying to find the reasons for the "movability" of vowels of the [æ]—[a] type from the diachronic point of view, the first step is to turn one's attention to the way these sounds are perceived. That this is a prerequisite to any attempt at understanding linguistic change has been amply demonstrated by Andersen (1973), whose model of abductive and deductive change we will follow here. The application of this model to the study of phonological change requires, among other things, that acoustic distinctive features be employed, rather than the more commonly used articulatory ones or any framework which includes these (e.g., the universal phonetic features proposed by Ladefoged 1971).

The motivation of this choice will become apparent when the following is taken into consideration. One of the main concerns of historical linguistics is to broaden our understanding of the origin of innovations which arise in the process of language acquisition (constructing a grammar of their language by language learners) and which, ultimately, lead to linguistic changes. The only data at the learner's disposal is what he hears, in other words — the output of his models. Consequently, the path for the linguist to follow is to try and get as close as possible to the way the learner perceives and interprets his data. Since our understanding of the perceptual or even auditory side of speech is not yet advanced enough, we have to content ourselves with acoustic parameters, which, at the moment, take us the nearest we can get to the destination of the message. The corresponding articulatory mechanisms involved in the production of this message, are one more step removed from it (cf. Jakobson et al. 1965:12f, Jakobson and Halle 1968:433 ff), which is why neither purely articulatory features nor any universal phonetic feature framework can meet our needs.

As far as the application of the abductive-deductive model to the present case is concerned, let it be stressed that — for reasons discussed later — we will be concerned exclusively with the primary abductive innovation in evolution change. This type of innovation arises when the output from which the learner infers his grammar is in some way ambiguous, i.e., allows of more than one interpretation. In the case of phonological changes the ambiguity may involve, among other things, different possibilities of bisecting the acoustic continuum of frequency. The bisect is necessary for the learner to discriminate between the grave and nongrave (acute) phonemes of his language, and, then, if needed, also between flat and nonflat (plain) or sharp and nonsharp (plain) phonemes.

Now, as was observed by Hyman (1973:336) and reported by Davidson-Nielsen and Orum (1978:205), it seems to be a phonetic universal that the higher a front vowel, the "more acute" it is, and the higher a back vowel, the "more grave" it is. In other words, although both [i] and [æ] are acute and both [u] and [a] are grave in the system

\[
\begin{array}{c|c}
[i] & [u] \\
[a] & [o] \\
[æ] & [a] \\
\end{array}
\]

the absolute difference in tonality is smaller between [æ] and [a] than between [i] and [u]. This means that [æ] and [a] are perceptually more alike than their counterparts in the two higher series. It seems legitimate then, to locate the source of the instability of WG a in English in the universal difficulty that learners of any language with more than one low vowel encounter in interpreting these as either grave or acute.

Having proposed this tentative answer to the first question put forward at the beginning, let us turn to our main problem and discuss the kind of innovation that may have been involved in the change of WG a from Late Old English into Early Middle English. Before proceeding to the analysis of the change itself, however, we have to specify what we believe the diachronic correspondence to have been, i.e., state the input to and the output of the change. This is by no means a trivial task, as there is no agreement among scholars either with regard to the nature of the opposition a/e in Old English or as regards the exact phonetic characteristics of ME [æ]. Needless to say, well-founded arguments for resolving the former issue, as well as a methodologically sound reconstruction in the case of the latter, are necessary conditions for any further considerations to be meaningful.

As far as the Late Old English state of affairs is concerned, we are not going to deal with Kentish and West Mercian, since in these dialects the situation was remarkably different than elsewhere, due to the early change [æ] > [e] in the former and to the so-called "second fronting" in the latter. For the remain-
ing dialects we take the situation to have been the following: there existed a phonemic contrast between /æ/ and /a:/ /æ/ having resulted from both Anglo-Frisian Brightening and the i-mutation of /a/. The originally allophonic opposition was phonemized some time in the 7th century, after i-mutation had run its course and certain changes in unstressed syllables had taken place. The contrast was weak, i.e., there were relatively few contrastive minimal pairs, but their number and the variety of their sources (especially in English) seem to guarantee the soundness of the biphonemic solution (see Chatman 1958; cf. also Ross 1951). Altogether, it will have been noticed that our stance concerning the character of the OE opposition /æ/ /a/ conforms to the “majority view.” The situation is rather different when it comes to deciding upon the most probable outcome of the change, i.e., upon the nature of ME /æ/, to the discussion of which we now proceed.

According to Jordan (1934: §§ 32, 56), the first occurrences of /æ/ for both /æ/ and /a/ of every source (i.e., also from /æ/ /æ/ and shortened /æ/ /æ/) appeared at the beginning of the 12th century. Until then, the traditional /æ/ (if normalized, also /æ/) had remained the predominant spelling, together with the occasional inverse /æ/ /æ/ /æ/. The final adoption of /æ/ has served historians of the English language as the basis for positing a merger of /æ/ /æ/ and /æ/, usually dated back to the period before or around 1100 (e.g., Lipkin 1914: § 363). It is, however, far from obvious whether the result of this change should be taken to have been a front or a back vowel. The more popular view is that ME /æ/ was back (e.g., Jespersen 1909, Jones 1950, Gimson 1965, Chomsky and Halle 1968), one of the few exceptions being Dobson (1968). A different interpretation (one postulating the existence of /æ/ in Middle English) is given in Stockwell 1959, Roszkiewicz 1951 and Eirik 1965.

Lass (1973) argues convincingly that the criteria made use of in proposing either the front or the back value for ME /æ/ are, on various grounds, questionable in all the above cases. Sometimes, as was the case with Jespersen and Dobson, the same arguments were used in support of the opposite solutions. The orthoepic evidence which has often served as the main source for reconstruction, namely the statement by Hart (1669: 30) that the vowel was articulated “with yde opening the mouth, as when a man yaseth,” is also of no value for deciding upon the frontness/backness of the sound in question. Finally, for Chomsky and Halle the choice of the back value was a direct consequence of their putative universal feature framework, which did not allow them to distinguish more than two nonround low vowels ([æ] and [a]), thus excluding [a] in principle.

According to Lass, the clue for reconstructing the position of ME /æ/ should be sought in comparative materials from modern English dialects. Because no major changes of WE /æ/ are reported to have taken place in the north of England since medieval times, the fact that there are no back reflexes of this vowel in any northern dialect at present forces Lass to conclude that ME /æ/ must have been a front vowel. His view has been fully supported by the results of Romaine’s work on Scots (Romaine 1982a).

Thus, if LOE /æ/ and /æ/ merged into EM/E /æ/, the primary abductive innovation involved in the change may have resulted from the fact that, due to its perceptual similarity, LOE /æ/ and /æ/ were both perceived by language learners as high in terms of tonality and, accordingly, reinterpreted as realizations of one phoneme, a nonround /æ/.

Of course, this is only a hypothesis as to what might have happened, not an explanation of why the innovation took the form it did (more specifically, why its result was not a grave vowel) or why it was, at a certain time, adopted by the community, thus resulting in a major qualitative change.

The model of abductive-deductive change provides the diachronic linguist with a kind of heuristic to be followed in establishing the operation of any linguistic change, starting with the initial abductive innovation and ending at the point where the change has gone to completion. Unfortunately, in trying to account for changes postulated for early stages of languages, the model can hardly be applied beyond the level of positing the primary abductive innovation. This is due to the scarcity of evidence at our disposal and to its exclusively written character. Therefore, although one could speculate on the theoretically possible adaptive rules, on the abductive innovations involved in their loss, etc., it would not broaden one’s understanding of the actual change, unless, of course, some additional data were found.

As observed by Coseriu (1974: 68), the initial innovation in the process of language change may have some sort of “physical” motivation, but its adoption by the community is definitely not a mechanical process. Being an act of choice and insight (albeit usually unconscious) on the part of the speakers, it can only be understood by reference to some cultural, aesthetic and/or functional goals which it serves. Since, for reasons given above, we have practically no clue to the possible cultural or aesthetic factors which may have favored the spread of the innovation in question (unlike in the case of changes of WE /æ/ in modern English dialects: cf. Labov 1969, Romaine 1982a), the only aspect of the change which we have some sort of access to remains the function it may have served in the language system, divorced from its social setting.

Incidentally, it will be evident from the start that this cannot result in any kind of causal explanation; the search for such explanations is, in any case,
misconceived with respect to changes in language (Coseriu 1974:152). When we claim that certain factors played a part in a given change, it should be kept in mind that, as Coseriu (1974:90) says:

"(...) es sich um passive Faktoren handelt, um Umschläge des Sprechens und historische Bestimmungen der Sprachfreiteit und nicht um aktive Faktoren ("Ursachen")".

Furthermore, the "passive factors" which will be postulated in connection with our change are only some of the factors involved.

Putting all things together, the interpretation proposed below is both non-causal and, necessarily, partial.

Having made all these reservations, let us proceed to the presentation of the solution which comes to mind in the case discussed. We believe that the functional aspect of the merger of LOE /æ/ and /a/ into ME /ə/ can be illuminated by viewing this change in the light of other developments in the system of short vowels, either prior to or roughly simultaneous with it.

The only major qualitative changes of stressed short vowels between Old English and Middle English were the loss of /y/ and the loss of /æ/. Campbell (1959: §§ 36, 317) places the beginnings of the unrounding of /y/ (both from i-mutated /i/ and from W-class) as early as 800. According to Luick (1914: §307), /y/ changed into /i/ in the North and the North-East Midlands in the course of the tenth and eleventh centuries, though the merger started to be reflected in writing only after the adoption of the grapheme (′y′) for /y/ in dialects where the latter still existed and after the consequent alternative use of (′i′) for /i/ in Early Middle English the process of unrounding was characteristic only of the above-mentioned areas and of a small part of the South-West (mainly Devonshire and Dorset). It probably reached London around 1300 and had been completed in the West Midlands, as well as in the remaining part of the South-West, by the end of the fourteenth century (Jordan 1934: §§ 39, 41).

As far as /æ/ is concerned, the few occurrences of this sound which had appeared as the result of the i-mutation of /i/ were in all dialects unrounded to /æ/ still in Old English (Campbell 1959: §§ 36, 196, 288). The unrounding of the /æ/ which developed from the monophthongization of /e/ took place in the East Midlands and the North in the course of the twelfth century. The West Midlands and parts of the South preserved the sound well into the fourteenth century (Jordan 1934: §§ 56, 65, 66).

Due to the fact that the rate of unrounding of both /y/ and /æ/ varied with locality, we have to exclude South-Western (West Saxon) from our considerations. As will have been recalled, we have already had to disregard Kentish and West Mercian (West Midlands), albeit for different reasons. Thus, in what follows we are concerned exclusively with the Northern and East Midlands dialects, which already in Early Middle English unambiguously show the loss of /y/ and /æ/, as well as the merger of /æ/ and /a/ into /ə/. Furthermore, for our purposes, the changes /y/ > /i/ and /æ/ > /e/ will be treated together and referred to simply as the unrounding of front rounded vowels.

It is now time to look at this unrounding in terms of the initial adverbial innovation involved in it, as well as in terms of its results for the whole system of stressed short vowels, and then to try and connect both these aspects with the merger of /æ/ and /a/.

In acoustic terms, /y/ and /æ/ are non-frontere flat vowels, the latter feature distinguishing them from /i/ and /e/ respectively. Thus, keeping in mind Andersen's (1973:768) definition of the phoneme as "a syntagm or construction whose ultimate constituents are phonemic 'feature values' and his claim that "the phonological terms which constitute a phoneme represent a hypothetic rather than a paraletic structure, in which any higher-ranking immediate constituent is of more central importance in the syntagm than the corresponding lower-ranking immediate constituent", we may conceive of the constituent which defines the tonality of the LOE front non-low short vowel phonemes as itself forming a hypothetic two-constituent syntagm, the higher-ranking constituent being the feature of gravity, which divides the continuum into high and low tonality, and the lower-ranking one - the feature of flattening, which further subdivides it into lowered and non-lowered.

What the loss of /y/ and /æ/ from the system amounts to is a simplification of this two-constituent syntagm:

\[
\begin{align*}
\text{LOE} & \quad \text{EME} \\
\text{flat} & \quad \text{non-lowered high} \\
\text{tonality} & \quad i \quad e \quad i \quad e \\
\text{lowed high} & \quad y \quad o \quad i \quad e \\
\text{tonality} & \quad \text{high} \\
\text{grave} & \quad \text{lowed high} \\
\end{align*}
\]

If we stop here to look for similarities between this innovation and the one involved in the merger of /æ/ and /a/, we can see that in both cases one has to do with the speakers' realigning of certain sounds (/y/, /æ/ and /a/) as "less grave", which means perceiving them as more high-pitched.

To bring out the parallel, we can juxtapose the two innovations in the following diagram:
The symbols between parentheses have been included in (4) for merely illustrative purposes. They represent phonemes which would have existed in the system in question, had the correlation of flattening in Old English involved not only noncompacted non grave vowels, but all the others as well. This does not seem to have been the case. In particular, contrary to Reszkiewicz (1971, 1972), we believe that there is no justification for postulating even a short-lived flat counterpart of /æ/ for any stage in the (pre-)history of English. As regards the flat equivalent of /a/, however, the situation seems less straightforward. Kuhn (1961:528) argues that the reflex of PG *a−nasal was phonemically (a) for a brief period in ninth-century Anglian. Hogg (1982), although admitting that there is insufficient evidence to support this suggestion for Mercian, puts forward rather convincing arguments in its favour with respect to Northumbrian. Even there, nonetheless, /a/ must have soon been eliminated through analogical change, giving the same phonemic distribution as in other dialects.

As may be concluded from (4), the structural innovations involved in both our changes appear to have served similar functions in the system, one of them making flattening nondistinctive, the other assuring that for each particular phoneme the feature flat has the same value as the feature grave. Diagrams (5) and (6), which present the stressed short vowel phonemes of Late Old English and Early Middle English respectively, provide another way of illustrating this:

<table>
<thead>
<tr>
<th>LOE</th>
<th>EME</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-lowered high tonality</td>
<td>i e æ</td>
</tr>
<tr>
<td>lowered high tonality</td>
<td>y ø (æ) i o a</td>
</tr>
<tr>
<td>non-lowered low tonality</td>
<td>(u) (y) a</td>
</tr>
<tr>
<td>lowered low tonality</td>
<td>u o (ø) u o</td>
</tr>
</tbody>
</table>

The only other case when flattening would in the same way agree with gravity in (6) would be a system in which the compact phoneme were grave and flat. One might ask then whether the greater likelihood of /a/ (rather than /æ/ or /ø/) as a result of the merger of /æ/ and /a/ is in principle predictable. There seems to be no doubt about this. If one looks at (4) again, it is evident that, in order to change into /æ/, /a/ would have to "travel" from one extreme of the tonality scale to the other (from nonlabeled high to lowered low tonality), which would be one step more than /æ/ covers in its shift to /æ/. Moreover, it appears most unlikely for two nonflat vowels to merge into a flat one, if only because spontaneous rounding, like spontaneous retraction, is certainly not typical of English.

Summing up: if flattening — which, having appeared as a result of i-mutation, became distinctive for the series of noncompacted non grave vowels and disappeared from it with the loss of /æ/ and /æ/ — was to become redundant in Early Middle English and to coincide with gravity in the way discussed, the only possibility for the merger of LOE /æ/ and /æ/ was a nongrave vowel, i.e., /æ/ or /æ/. It should be noted at this point that, if one is interested only in flattening becoming redundant, the result of this change is irrelevant, the merger itself producing the effect. Such a view is presented, e.g., in Endmann (1972: 46ff).

The above conclusions are, of course, not tantamount to saying that phonological changes can in some way anticipate their own effects and modify their direction accordingly. Normally, there can be no a priori goal to a change. As argued by Coseriu (1974:194), each change is the result of numerous consultations in the same direction. It seems likely then that those innovations whose effects may eventually serve some (however broadly defined) purpose will be adopted more readily than others.

In view of the rather controversial status of teleological explanations in historical linguistics, we should perhaps comment in greater detail on the perspective adopted here. Since the pioneering work of Jakobson (1928, 1931, 1932), the notion of teleology in general and the legitimacy of functional explanations in particular have frequently been attacked. One of the most serious blows came recently from Lass (1980). The reaction to Lass’ hook, however,
seems to suggest that teleological explanations of some sort are still commonly regarded as providing the most important type of answer to the HOW-question of language change. Since the presentation of the arguments adduced in their favour is beyond the scope of the present paper, we can only refer the reader to Itkonen (1978 and 1981) and to the papers from the Round Table on Diachrony (Lille 1982) mentioned below.

As has already been signalled, the crucial problem is that explanations by reference to the (final effects in terms of) function of a change are HOW- and not WHY-explanations. The latter are, in any event, not needed in the case of (unconsciously) rational actions, to which language behaviour seems to belong. According to the "principle of unconscious rationality", a single language change is a collective action, whose acceptance by the community functions as a 'rationality filter' on innovations (Itkonen 1981:693).

More specifically, as argued by Harris (1982:8), "speakers will, in general, prefer - unconsciously, of course - to organize the material which comprises their language in a maximally regular and economic fashion, a preference which manifests itself in the changes captured by the various generalizations and taxonomies diversely formulated by diachronic linguists of various persuasions for many decades, from 'structural pressure', through 'rule simplification to 'typological consistency'."

To return to the specific problem discussed in this paper, we have argued that the loss of front rounded vowels and its consequences for the vocable system of Middle English may have had some bearing both on the coalescence of /æ/ and /æ/ and on the output of this change. More generally, it could be claimed that the loss (or rather the subsequent losses) of /e/ and /æ/ on the one hand, and the merger of /æ/ and /æ/ into /æ/ on the other, constituted a kind of feedback, in the sense that when one of those changes had taken place, the likelihood of the occurrence of the other was increased. Formulated in this way, the conclusion would be altogether independent of the relative chronology of the developments in question and could perhaps be extended to other ME dialects, apart from those discussed.9

As regards the implications of such an approach for the study of WG a in English, the present author believes that - if the hypothesis formulated here is considered of interest for the understanding of the change in question - one should not neglect the possibility of examining other developments in the complicated history of this vowel along similar lines. Since the general reasons behind the instability of WG a seem to be quite obvious, what remains to be done is to establish the conditions which made possible the particular changes that took place in it at different times and in different dialects. One place to look for these — especially when early stages of the language are concerned — is in the phonological structure of the language and all its modifications which were contemporary with the one(s) we want to account for.

REFERENCES


---

9 Cf. e.g., Sundby (1963:196, 200) who discusses the situation in the South-Western and West Midlands dialects (where the chronology of the changes appears to have been the reverse of the one agreed upon for the dialects dealt with here) and notices a systematic connection (expressed in terms of articulatory space) between the merger of /æ/ and /æ/ and the subsequent loss of /æ/.

---