Vowel Neutralization and Vowel Reduction

Patricia Donegan Miller
Ohio State University

In the study of vowel systems—perhaps more than in the study of consonants—two dominant and opposing forces in phonology become especially apparent: the principle of maximal contrast, and the principle of least effort. This second principle, that of least effort, is represented in a particularly clear way in a process that I will call Vowel Neutralization. In proposing this process, I am going to attempt to characterize as due to a single innate or natural process 1.) the severe limitations on certain very small vowel systems (including those of child language), and 2.) the fairly common phenomenon of vowel reduction. In adult languages, application of neutralization collapses all the vowels of the target language to a neutral or schwa-like vowels for more distinctive ones. Besides being especially favored for unstressed non-tense vowels, Neutralization may be hierarchically conditioned by the quality of the vowels to which it applies, or it may be limited by context. Since it is only in vowel reduction, however, that these limitations occur, they will be examined in the part of this paper that deals specifically with vowel reduction. Since the process in its fullest form involves the substitution of a neutral vowel for any other vowel, it is easy to see how it could help account for the child's substitution of a for all adult vowels. The vowels the child hears are replaced by or schwa. Then another process operates, which lowers stressed non-colored vowels to , and is the child's output, since his vowels are, at this point, all stressed.

When the child learns to suppress a part of the Neutralization process (usually he learns to stop making all vowels non-high), he acquires an opposition. Then his rule would look like this:

\[
V \rightarrow [-Round] \\
[-Palatal] \\
[-Low]
\]

When lowering operates, it may then lower only the non-high neutral vowel, leaving the child with and .

I am aware of one case where this remains the child's system: is substituted for high vowels and for mid vowels with the exception of and replaces low vowels and (Oller forthcoming). The evidence that the child's arises from a mid vowel's being stressed is scant, but something does exist: the only word in his corpus with an unstressed initial syllable, "banana," was consistently pronounced .

Ordinarily, however, instead of retaining an in the two-vowel stage, children will either round or palatalize the high vowel of . In this way, Joan Velten (Velten 1962, 25 et passim) had an system, and Hildegard Leopold (Leopold 1953-54, 353 at passim) had and .

There are adult vowel systems, too, that display the effects of Neutralization. In certain languages, timbre or coloring seems not to be distinctive; only height distinctions are made. An example of such a language would be Kabardian. Although Trubetzkoy interprets the vowel system of this language as having three heights or degrees of sonority, it can be interpreted as having only two vowels, and ; or, if the occurrence of is predictable, perhaps it has only one, . If the language is regarded as having vowels at all, will represent the high vowels, and the mid and low ones, with the relevant distinction being one of open vs. close (Kutepes 1960, 23 et passim).

Such a system can readily be generated by suppressing only the height-removing function of the Neutralization process. Note that the speakers would also have to suppress the processes that round or front non-low colorless vowels and that they must limit the lowering process.

This, of course, makes the prediction that speakers of such languages, given "foreign" sounds, will substitute vowels of the corresponding height, but not necessarily of the corresponding timbre. (Timbre might be determined by the surrounding consonants, but that is not an immediate concern here.) The loan phonologies of such languages would reflect such substitutions, but I do not know of any work on this subject.
In these languages with vertical vowel systems, the centralization accomplished by the Neutralization process may be undone by contextual variation, but since these variants are not distinctive, it seems justifiable to attribute the lack of distinctions to a process operating at some higher-than-phonetic level.\(^3\)

When the application of Neutralization is limited to unstressed vowels, vowel reduction occurs. I should point out here that—as the very use of the term "unstressed" implies—I am inclined to believe that only stress languages (and not pitch accent languages) exhibit vowel reduction of the sort that I am talking about here—i.e., reduction that affects vowel quality.

But whether vowel reduction is limited to stress languages or not, this process does exhibit certain differences from its "mother-process," Vowel Neutralization.

First, when neutralization is sufficiently limited so that it applies only to vowels without stress, it is usually further limited so that it removes only some of the positive feature specifications of the vowels. To be more positive about this syntactically, reduced vowels in some languages (or in some contexts) may retain certain distinguishing features. That is, they may not become completely neutral.

For example, if the speakers limit the process

\[
\begin{align*}
&V \rightarrow \begin{cases}
&\text{[-Stress]} \\
&\text{[-High]}
&\text{[-Low]}
&\text{[-Palatal]}
&\text{[-Round]}
\end{cases}
\end{align*}
\]

then the reduced vowels will retain their original rounding, and the language will have a reduced-vowel system consisting of a round and a non-round vowel. In this way, a system like that of Eastern Ojibwa can be achieved, where \(a\) is the reduced variant of the lax vowels \(a\) and \(u\), and a round indeterminate vowel, like a rounded schwa, \(o\), is the reduced variant of lax \(o\) (Bloomfield 1956, 5–6).

Some dialects of English have an example of the retention of the \(+\) Palatal specification on reduced vowels, although this retention occurs only in specialized environments: before high consonants, the original palatality of the vowel is maintained. This accounts for the distinction between epoch and epic and between the final syllables of comic and havoc. Note also the pairs manic/maniacal, and courageous/courage, in which the reduced vowel of the second member is \(\ddot{u}\). Here Neutralization is limited to something like

\[
\begin{align*}
&V \rightarrow \begin{cases}
&\text{[-Stress]} \\
&\text{[-High]}
&\text{[-Low]}
&\text{[-Palatal]}
&\text{[-Round]}
\end{cases}
\end{align*}
\]

Since it is possible, then, to limit reduction in such a way as to maintain rounding or palatalization, one would expect to be able to retain a height distinction in a reduced-vowel system in a similar way—i.e., to have a reduced system with two vowels, \(i\) and \(\ddot{a}\), or \(\ddot{a}\) and \(\ddot{e}\). One would expect this all the more since the removal of height distinctions is the first part of the process suppressed by children, and since it may be the only part suppressed in those languages with "vertical" vowel systems. That is, height distinctions are the first ones acquired and sometimes the only ones maintained.

Oddly enough, reduced-vowel systems do not seem to maintain height distinctions. (At least, I do not know about any height-maintaining systems.) On the other hand, though, perhaps this lack is not too odd: unlike the full, context-free process of vowel neutralization, vowel reduction is characteristic of unstressed vowels, especially in fast speech.

Why height should be more susceptible to Neutralization in unstressed vowels and color more susceptible in stressed vowels, I cannot say for sure, but a physiological explanation suggests itself. If it is the case that lack of stress is accompanied by lack of articulatory effort, and if part of this lack of effort is failure to increase jaw opening for vowels that are originally more sonorant, than a failure to make distinctions of vowel height in reduced positions does not seem extremely peculiar.

Perhaps the increase in jaw opening might be a grosser movement which is more easily acquired but which, since it requires more effort, is more likely to be abandoned.

There is another somewhat related respect in which vowel reduction appears to differ markedly from Vowel Neutralization, but I think that this apparent difference is not a difference in the Neutralization process itself, but is rather a difference in the way subsequent processes affect the stressed vowels that result from Neutralization and the unstressed vowels that have been neutralized in the reduction sense.

The difference is that stressed uncolored vowels lower and unstressed uncolored vowels raise. Thus the child's stressed \(\dddot{a} \rightarrow \dddot{a}\), while the adult's unstressed \(a \rightarrow \dddot{a}\) (\(\rightarrow \dddot{a}\):

\[
\text{"He had gone." } \rightarrow [\dddot{hi} \dddot{ad} \dddot{g\dddot{on}}] \rightarrow [\dddot{hi} \dddot{ad} \dddot{g\dddot{on}}].
\]

This raising of unstressed uncolored vowels may be due to the same tendency to fail to increase jaw opening that might account for the lack of height contrasts in reduced vowels.

In parallel to this suggestion and, in some sense supporting it, is the hierarchy of applicability of vowel reduction. When only some of the vowels of a system are reduced, the reduction often seems to proceed in an order from the low to the high vowels. In Breton, for instance, \(a\), \(o\), and \(\ddot{a}\) are reduced in moderately fast speech, and \(\ddot{a}\), \(\ddot{e}\), and \(\ddot{u}\) only in faster speech; \(i\), \(u\), and \(\dddot{u}\) apparently do not ever reduce enough to lose their character as high, mutually distinguishable vowels (Dressler 1972, 30).
Height does not seem to be the only factor that inhibits reduction, however. Rounding, too, often limits the scope of this process. In Modern Hebrew, for instance, lax $e$ and $a$, and sometimes $i$ will reduce to $e$ in fast speech. $i$ reduces only after $e$ and a $d$--i.e., only in very fast speech--but laxed $o$ ($u$) and $u$ do not appear to reduce at all (Semiloff-Zelasko, forthcoming). Here the order of reduction is not just low-to-high; it is more like low-to-high cross-cut by the Round specification of the vowels to be reduced. The rule might be written to look like this:

$$\begin{array}{c|c|c}
\text{V} & \text{High} & \text{Low} \\
\hline
\text{-Stress} & - & - \\
\text{-Tense} & - & - \\
\text{-Round} & - & - \\
\text{lower} & & \\
\end{array}$$

It sometimes appears that there are other factors that limit or encourage vowel reduction, too, such as syllable closure, or the position of the reducing syllable within the word.

In Latin, height, rounding, and syllable closure combine to constrain the special set of vowel reductions found in medial syllables in the period of probable initial stress. The non-high vowels $e$, $a$, and sometimes $o$ reduce to $e$ in unstressed (medial) position, and $o$ when it appears in an open syllable (Buck 1933, 100). Note that the reduction of the round vowel is more restricted than that of the non-round ones, and that closed syllables tend to inhibit this reduction while open syllables allow it. The further reduction of this $<e>$ to $<i>$ in open syllables only reaffirms the role of syllable closure as an inhibitor of reduction.

This secondary reduction of vowels in open syllables to $<i>$ or sometimes to $<u>$ may reflect to some extent on the quality of the reduced vowels as they may have been pronounced. The fact that, for some items, there was at one time a fluctuation between $<i>$ and $<u>$ would indicate that the actual reduced vowel was something like an $\hat{e}$, the raised variant of the schwa which orthographically $<e>$ could represent.

Thus, the reduction process in Latin medial syllables could be described by saying that $e$, $a$, and sometimes $o$ were first reduced to a schwa-like vowel (written $<e>$), and were then further reduced, in open syllables, to $\hat{e}$ ($<i>$ or $<u>$). Then, when the stress shift occurred, these centralized vowels were given the quality of full vowels--$e$ for $\hat{e}$ and usually $\hat{i}$ for $\hat{u}$, but sometimes $u$ for $\hat{u}$, when there were labials in the environment).

There are, of course, other examples of vowel reduction where the final output is not a mid central vowel. A well-known example is the akanye and ikan'ye of Modern Russian. As I understand the most basic facts of these phenomena (for Moscow Russian), the reductions affect only non-high unstressed vowels:

$\hat{e}$ reduces to $o$ or $\hat{a}$ in pretonic or absolute initial positions and to $\breve{a}$ elsewhere; $\hat{a}$ reduces to $i$; and $\hat{e}$ remains $\hat{a}$ or $\hat{e}$ in postonic consonants, and it reduces to $\breve{a}$ elsewhere (Kondrasov 1962, 85).

In terms of a vowel reduction process based on Neutralization, these occurrences may be looked at somewhat differently. If the $e$, $a$, and $o$ (except $o$ after palatalized consonants) to $\hat{e}$, then $\hat{a}$ is slightly to $i$. A subsequent lowering of the remaining schwas to $o$ in an initial and pretonic positions would provide the proper reduced-vowel distribution.

The reason for my use of the Latin and Russian examples is that they raise two rather puzzling questions regarding the environments in which such reductions occur. The first of these two regards Latin: why is it that the open medial syllables apparently favor reduction and closed syllables disfavor it?

In answer to this, I would like to suggest that there may have been, during Latin's initial-stress period, some sort of minor stress on closed medial syllables. This suggestion is based on an analogy with certain "half-stressed" syllables of Old English, which also had a basically initial stress. According to Campbell (1939, 34-35), "any long final syllable" has been described as a syllable that such a syllable must be long by virtue of having two consonants after the $a$ vowel--i.e., that it must be closed, after another long syllable or its equivalent, acquires half-stress when it becomes internal by the addition of an inflexion: "Hengestes, æhwelclina. Short final syllables which become long in inflexion are similarly treated: òperne." Thus, closed medial syllables in another initial-stress language acquire half-stress.

If Latin closed medial syllables had a minor stress like this, it may have been enough to prevent 1.) the reduction of $o$ in such syllables, and 2.) the second step in Latin vowel reduction--the reduction of $<e>$ to $<i>$ above. Such a suggestion would also explain the sensitivity of the reduction rule to syllable closure, and it would indicate that the sensitivity was really to degrees of stress.

The second question--that regarding Russian--has to do with why $\hat{a}$ should become $a$ in pretonic or absolute initial position. A suggestion in answer to this question rests on the view of Russian as a rising-accent or stress anticipation language. A similar case would be Modern Hebrew, where $a \rightarrow \hat{\alpha}$ two syllables before the stress, but not immediately before it. Such a view would account for the pretonic lowering, but it would not explain lowering in initial position. Of course, some type of stress difference in this position could be posited, but as far as I know, this is entirely speculative at this point. The fact remains, however, that vowels in this position act as if they have some element of stress, so a stress element in this position is something to look for.
What I have tried to show in this paper is that, despite certain differences, wholesale vowel neutralization and various vowel reductions can be traced to the same innate process, given a framework of natural phonology. The common factors of centralization, removal of rounding and palatalization, loss of height, and loss of sonority suggest that vowel reduction is a limitation or subprocess, sensitive to factors of vowel quality and especially to stress, of Vowel Neutralization.

Footnotes

1. The fact that the mid vowels for all of these children group with the high vowels appears to be due to the operation of other processes before the operation of Neutralization. Briefly, the order of the processes relevant here would be:
   a. removal of color from low vowels (e, 3, etc. → a)
   b. raising of colored vowels (e, o, û → i, u, ü)
   c. neutralization
   d. colorless vowel lowering (of the mid vowel only, here).
In the case of adult systems, the ordering may be different. Reordering vs. Suppression is a question I have not yet dealt with in any detail.

2. I am also aware of Kuipers' vowelless interpretation, which views openness of release as a feature, but I do not think the distinction is relevant here.

3. In Stampe's present terminology, the neutralization process would be a dominating, context-free one, and the subsequent context-sensitive, color-adding processes would be dominated.

4. I do not insist, of course, that <e> and <i> were necessarily a and â. It certainly seems to me to be a reasonable assumption, but it is not intended as support for the process I am discussing. What I intend to suggest is that Latin "vowel weakening" is a real reduction, compatible with the neutralization process, because I find the conditions under which this reduction occurred interesting, and I would expect to find other examples of vowel reduction constrained by syllable closure.

Bibliography

Stampe, D. Forthcoming. Natural Phonology.
PAPERS FROM THE
EIGHTH REGIONAL MEETING
CHICAGO LINGUISTIC SOCIETY

APRIL 14-16, 1972

Edited by:
Paul M. Peranteau
Judith N. Levi
Gloria C. Phares

Chicago Linguistic Society
Foster Hall
1130 East 59th Street
Chicago, Illinois
60637