Synthetic English Speech by Rule

M. Douglas McIlroy
Bell Telephone Laboratories, Incorporated
Murray Hill, New Jersey

ABSTRACT

A compact program produces synthetic speech from English text via an inexpensive commercial synthesizer. The speech is not inflected, but within that constraint is deemed adequate (that is, about as intelligible as can be obtained using the device) on at least 97% of running text. The program works almost entirely by rule, most of which have a stereotyped form. Since no large dictionary is needed, secondary storage is not used. The entire body of synthesis rules is given, together with examples of their behavior.

Neither the program nor the output device pretend to produce natural speech, but it is intelligible to everybody who has heard much of it. Besides its modest cost and size, the prime virtues of the program are its accessibility and its readiness to speak anything, albeit far-fetchedly on occasion. The scheme should also prove useful as a method of last resort in concert with more sophisticated synthesis procedures.

March 1974

Synthetic English Speech by Rule
Addenda and Errata
April 11,1974

Replace 4.1(b) by

(b)	[.,;:?!'"])]*#	*—0 → #*

After 4.4(b') move *tenebrious* from "Examples" to "Exceptions".

In 4.5(b) and 4.5(c) replace [*aeiouwxy*] with [*aehiouwxy*], and add *dihedral* under "Compare".

Delete *eye* from list of Appendix 1.

Before 7.16.2 of Appendix 2 insert

7.16.1½	*#eye*	*→I*	*eyelash*	*pinkeye* *greyer*

Between 8.2.2.3 and 8.2.2.4 of Appendix 2 insert

8.2.2.3½	*#gil*	*g—il*	*gilbert* *gild*	*gillette*	*fragility*

Replace 17.10.4 of Appendix 2 by

17.10.4	*foot*	*f.oo—t*	*hotfoote|d*

In 22.7 of Appendix 2 replace phonetics by *sh,sh.*

Make changes in Appendix 2A corresponding to those in Appendix 2.

Add to Appendix 3 the following correspondences with the manufacturer's mnemonics.

a0	AH	*er*	ER	*ie*	IE	*yu*	U
a1	AH1	*eu*	OOH	*ih*	IH	*iu*	U1
aw	AW	*eh*	EHH	*o0*	O	*ju*	IU
au	AW1	*y0*	EE	*o1*	O1	*sh*	SH
ae	AE	*y1*	Y	*o2*	O2	*zh*	ZH
ea	AE1	*y2*	Y1	*ou*	OO1	*j*	J
ai	A	*ay*	AY	*oo*	OO	*ch*	CH
aj	A1	*i0*	I	*u0*	UH	*th*	TH
e0	EH	*i1*	I1	*u1*	UH1	*dh*	THV
e1	EH1	*i2*	I2	*u2*	UH2	*ng*	NG
e2	EH2	*iy*	IY	*u3*	UH3		

In Appendix 4 move *eyelash* from "Bad words" to "Good words".

Synthetic English Speech by Rule

M. Douglas McIlroy
Bell Telephone Laboratories, Incorporated
Murray Hill, New Jersey

1 Introduction

The Computer Science Center at this laboratory has experimented with an inexpensive speech synthesizer [1] as a regular output device in a general purpose computing system. Our intention was not to do speech research or to create artificial speech as an end in itself. In the present state of the art[2,3], those goals require much more elaborate facilities than we have at our disposal.

We wished to see what uses might evolve when speech became available more or less on a par with printed output. For this goal, "naturalness" was not a prerequisite, any more than it is for printed output. Most computers still print mainly in upper case, are incapable of printing mathematical notation, and normally produce cryptic codes or tabular stuff that require considerable indulgence to be understood. Since printed gobbledygook is so widely accepted from computers—and fed into them, witness any manufacturer's operating system manual—we suspected that spoken gobbledygook might be quite passable, too, except for one severe difficulty: Being ephemeral, sounds must be understood at first hearing. As it turns out, long speeches *are* hard to understand, as are extremely short utterances of very simple words out of context. But given a little familiarity with the machine's "accent", one finds short sentences to be quite intelligible.

Since we hoped to use the voice synthesizer as a regular output device for other programs, the speaking routine had to be small enough and fast enough to coexist with other programs that do the actual computing of interest. Needless to say it had to produce connected speech. Furthermore it ought to be ready to speak most anything with relatively little coaxing, and be tunable in considerable detail when special vocabularies (e.g. RxPch) or the inevitable pronunciation gaffes did arise.

2 Evolution

The detailed properties of our present speaking program were not even vaguely anticipated when the voice synthesizer was installed. Once it was in the operating system as a full-fledged IO device, we used a program for driving it from phonetic strings. We created files of single utterances, ranging from household cries like "lunchtime!" to classical quotations, including the inevitable soliloquy of Hamlet. Tuning was painful, so we quickly built an associative program that could store, update and print the whole vocabulary or single words, and also could pronounce any stream of words from the vocabulary. Words that it didn't know were split up into constituent one-letter words; in effect they were spelled out. Various useful vocabulary fragments developed quickly—the alphabet and other characters, the days and months, the 400 most common words of English, names of visitors to the lab ("welcome to UNIX, Joe"), chess moves, etc. A few hundred man-hours so expended gave us a good feel for phonetic coding, but didn't really go far to give a generally useful output device. Experience did show, however, that for this particular machine some short cuts were possible. In particular we soon gave up on inflection almost completely. It was clear that much of the encoding had become cut-and-dried; at least half of even the most common words of English, among which irregularities abound, were trivial. So I was led to upgrade the successful associative learning program to accept pronunciation rules for word fragments as well as words. I also built in a few more complicated rules such as the final silent *e* paradigm.

The last-mentioned program is still in use. Its tables now contain about 750 rules, including about

100 words
580 word fragments
70 letters

The rules occupy about 11000 bytes on a PDP11/45. Written in a higher level language with little attention to efficiency, the program runs at about 15 words per second of CPU time. There are about 4500 bytes of pronouncing code, including table search and the special hand-coded paradigms, plus 1900 bytes of code for interactive display and maintenance of the tables. The program is completely self-contained; no other supporting programs are needed, nor are any used. The program can be driven by typing into it, or by "piping" the output of any other process on the machine into it in real time[4].

The program breaks the input up into "words", delimited by spaces, certain punctuation characters and line breaks. The pronunciation of each word is found by a succession of increasingly desperate trials, which will be elaborated in Section 4:

(1) See if the whole word is in the table given in Appendix 1.
(2) Map capitals into small letters, strip punctuation, and try step 1 again.
(3) If the word ends in *s*, strip the *s*. Change final *ie* to *y* regardless of whether there was a final *s*. If any changes were made, try step 1 again.
(4) Try automatic pronunciation: be elaborated further on:
  (a) Reject a word consisting of 1 letter or a word without a vowel.
  (b) Mark some common endings, including *e*, and mark long vowels indicated by final *e* or equivalent endings such as *ed* or *able*.
  (c) Mark potential long vowels such as *u* in *modular*, *e* in *aphelion*, and *a* in *outrageous*.
  (d) Mark probable medial silent *e* as in *bumblebee* and *minesweeper*, and mark long vowels indicated thereby.
  (e) Mark potential voiced medial *s* as in *visible* or *prismatic*.
  (f) After these marking steps, replace any stripped final *s*; then go through the word from left to right applying pronunciation rules to word fragments.
(5) When all else fails spell the word, punctuation and all. On letters for which no spelling rule exists, emit a burp.

The rules for whole words (1), word fragments (4f) and spelling (5) are dynamically alterable, all other steps are hand coded. The distinction among the three kinds of variable rule is important. For example there is a rule of each kind for *a*, each having a different pronunciation as shown here in vs (voice synthesizer) phonetics. These phonetics are defined in the appendix. The phonetics themselves will be of little use except in connection with the one particular synthesizer, but the left hand sides, and the rewrite portions of the rules, should be widely applicable in generating cheap approximate speech.

word *a*	*u3,u2,u3*
fragment *a*	*ae*
letter *a*	*ai,ay*

3 Notation

In the sequel brackets indicate a choice among a set of letters. For example [*aeiouy*] means any vowel. Boldface letters in sets stand for both upper and lower case; thus [**aeo**] means the same as [*AEOaeo*]. The mark # is placed at both beginning and end of every word. The sign ¬ complements the set, so ¬[*aeiouy*] means any nonvowel (including #). The notation [...]* means a string of zero or more letters from the bracketed set.

A vertical bar | marks potential silent *e* positions. Upper case letters denote potential long vowels or voiced *th* or *s*. Rewriting rules are shown with an arrow as in

ea	→ *E*
#sch	→ *sk*

Ellipses may be used on the right hand side of rewrite rules to make them shorter. In such cases underlines are used on the left to show parts that are rewritten; all other parts of the left hand side are understood to appear unchanged on the right. A bracketed set of characters on the right always stands for a copy of the string that matched the corresponding set on the left. For example this rule describing the handling of silent *p* in words such as *pneumatic* and *psychology* stands for both the rules that follow it.

#p[*ns*]	→ [*ns*]
#pn	→ *n*
#ps	→ *s*

Pronunciations are given by strings of vs phonemes separated by commas preceding the rewrite part of rules:

s#	*z*
E	*ee*
tio	*sh,ih,u2*
al|	*u2* → *l*
cia	*sh,ih* → *a*

With minor syntactic changes, rules of the above forms are understood by the program. Rules involving character classes however must be broken out into an equivalent set of rules for the program. They are used here for brevity.

Examples are given with each rule. Examples are shown exactly as processed, with rewriting taken into account; but exceptions and comparison words are written straight except in unusually obscure cases. When classes of exceptional words exist, they are illustrated by one or two words; these should be readily distinguishable from freakish exceptions, which are listed with equal billing. An additional list of words for comparison shows nearby words that are not covered by particular rules, or words intended to suggest why plausible alternative rules have not been used.

4 Preprocessing Rules

Most of the rules in this section involve rewriting only without pronunciation. To save space they are often stylized in a way that eliminates the possibility of expressing pronunciation. Examples are shown exactly as processed by the program with rewriting taken into account; but exceptions and comparison words are written straight, except in unusually obscure cases.

4.1 Capitals and Punctuation

When a word is not found verbatim in the vocabulary, it is standardized by stripping off leading and trailing punctuation and mapping an initial upper case letter into lower. The leading punctuation characters are indicated in the following rule. In this rule and the next only, the literal signs [ ] are quoted by making them bold.

(a) `#[([{\`"]*` → `#`

A hyphen that cuts off a word of 2 or more characters after leading punctuation, if any, is treated as a blank, except that it is restored if the spelling step (5) of the previous section becomes necessary. Trailing punctuation causes a pause in the speech. The trailing punctuation characters are:

(b) `]..,:;?!'"[)]*  −0→ #`

In lines where every letter is upper case, all are mapped to lower. Single words in upper case are not so treated, so that abbreviations like *UK* will be spelled out. As a result pronounceable acronyms of the kind so popular in computing, e.g. *FORTRAN*, also get spelled.

4.2 Final S

Final *s* is stripped, together with a preceding apostrophe, if any. Then, regardless of whether anything was stripped a final *ie* is rewritten as a *y*.

(a) ¬[*us*]*s*# → ¬[*us*]#
(b) '# → #
(c) *ie*# → *y*#

If the resulting word is in the vocabulary it is pronounced with a voiced or unvoiced *s* according to whether it does not or does end with an unvoiced consonant. In effect

(d) ¬[*cfkpt*]*s*# → ...*z*...

Though rule (a) causes some unnatural stripping of final *s*'s, it is usually harmless, for this stripping rarely leads to further mistreatment.

Examples	Exceptions	Compare
fits → *fit*	*yes*	*menus*
minus → *minus*		
lens → *len*		
series → *sery*		
pie → *py*		
laddie → *laddy*		

4.3 Final E

When final *e* is the only vowel in the word it is generally pronounced long. Otherwise final *e* is usually silent.

(a) #¬[*aeiouy*]*e*# → ...*E*...

Examples	Exceptions
me → *mE*	*re*
she → *shE*	*the*

Final *e* and a number of equivalent suffixes preceded by a vowel then a single consonant generally indicate that the vowel is long. The following endings are located and marked repetitively, provided a vowel remains in the rest of the word. The mark | is placed just before the first letter of the suffix, or just after that letter if it is an *e*, as indicated below.

|*able*	*e*|*ry*	|*ly*
|*ably*	*e*|*st*	|*ment*
e|*d*	*e*|*y*	|*ness*
e|*n*	|*ing*	|*or*
e|*r*	|*less*	

If the remainder ends with one of the following two endings, the ending is marked (under the same proviso that a vowel remain) and final *e* processing ends:

|*ic*	|*ical*

Lastly *e* itself is located and marked under the same proviso. However *e* before *e*, as in *indeed*, is not marked, and terminates final *e* processing. The marks, which typically indicate potential silent *e*'s, may be placed far inside of words.

Examples	Exceptions
arrangement → *arrange*|*ment*	*apostrophe*
forcefully → *force*|*ful*|*ly*	*implement*
sparseness → *sparse*|*ness*	
topical → *top*|*ical*	

If the leftmost ending begins with [*aeio*] then one of rules (b) and (c) may be applicable, subject to the exceptions expressed by rules (d) to (g) below. Some of the rules cause potential long vowels to be marked by capitalizing them—a convenient practice, but one that wreaks havoc with words like *McIlroy* that contain real capitals.

(b) ¬[*aeo*][*aeiouy*]¬[*aeiouywx*] ending → ...[*AEIOUY*]...

(c) ¬[*aeo*][*aeiouy*]*th* ending → ...[*AEIOUY*]*TH*...

(c') [*aeo*][*aeiouy*]*th* ending → ...*TH*...

Examples	Exceptions	Compare
bather → *bATHe*|*r*	*hyperbole*	*house*
sincerely → *sincEre*|*ly*	*derivative*	*sewer*
bite → *bIte*|	*infinite*	*growing*
potable → *pOt*|*able*	*gone*	*create*
vacuole → *vacuOle*|	*risen*	*dealer*
triune → *triUne*|		*axe*
style → *stYle*|		*boone*
strIated → *strIAte*|*d*		*trailing*
spitefulness → *spIte*|*ful*|*ness*		
breathe → *breaTHe*|		

There are several special cases. First, monosyllables whose only vowel is final *y*, *ie* or *ue* and their derivatives are recognized. In this rule 'ending' may be empty.

(d) #¬[*aeiouy*]*[*iuy*] ending → ...[*IUY*]...

Examples	Exceptions	Compare
cried → *crIe*|*d*	*priest*	*value*
fly → *flY*	*pier*	
driest → *drIe*|*st*	*ski*	
slyly → *slY*|*ly*		
blue → *blUe*		

Second, derivatives of certain words, usually polysyllabic, ending in *ce* or *ge*.

(e) [*aeiouy*]¬[*aeiouy*]*[*aeiouy*][*cg*] ending → ...

Examples	Exceptions	Compare
menaced → menace|d	*misplace*	*lice*
carriage → carriage|	*outrage*	*stage*
college → college|	*osage*	*mortise*
siege → siege|	*device*	*advise*
service → service|		

Third, derivatives of words ending in *el*

(f) *el* ending → ...

Examples	Exceptions
corbeled → corbele|d	*dele*
traveling → travel|ing	*allele*
celery → cele|ry	

Certain endings are equivalent to final *e* in disyllables where the only other vowel in the word is the indicated long vowel. Although placed here for explanatory purposes, rule (g) must be applied before rule (b) so that *able* in *stable* may be treated differently from *able* in *capable.*

(g) #¬[*aeiouy*]*[*aeiouy*]¬[*aeiouywx*] ending → ...[*AEIOUY*]...

The pertinent endings are listed below with the examples. They are not marked when found.

al *le* *re* *us* *y*

Examples	Exceptions	Compare
cabal → cAbal	*canal*	*lethal*
spiral → spIral	*metal*	*mothy*
able → Able	*treble*	*democracy*
bugle → bUgle	*grille*	*inimical*
cycle → cYcle	*barre*	*capable*
ogre → Ogre	*genre*	*isle*
lucre → lUcre	*latus*	*wiseacre*
locus → lOcus	*many*	*showy*
stylus → stYlus	*city*	*axle*
spiny → spIny	*copy*	

Some families of final *e* words still defy rule-making. One family is words ending in *ine*, which have several widely different pronunciations. Another is words ending in *is*. Singular words in *is* should be treated like the *us* words in rule (f) above, but this treatment breaks down completely for plurals of words ending in *i*. Some examples of these classes:

brigantine	*iris*
gelatine	*alibis*
valentine	*taxis*
aborigine	*otis*

4.4 Long Medial Vowels

For monomorphemic words—words that are not decomposable into affixes or multiple stems—a well-known rule states: An isolated vowel, followed by a single consonant then a vowel is pronounced long[5]. (From the standpoint of pronunciation, *x* is not a single consonant.)

Examples	Exceptions	Compare
major → mAjor	*camel*	*vixen*
relay → rElay	*metal*	*parameter*
minus → mInus	*digit*	*molecule*
fugue → fUgue	*model*	

As the comparison words indicate, this rule does not generally apply to polymorphemic words, and so is useless in its baldes form. However, it does apply reliably to *u*. The role of the separating consonant may also be played by certain consonant-*r* combinations, as indicated by the alternate rule in braces.

(a) ¬[*aeiou*]*u*{¬[*aeiouwxy*|]|[*bcdfgkpt*]*r*}[**aeiouy**] → ...*U*...

Examples	Exceptions	Compare
tumultuous → tUmultuous	*cutoff*	*uncle*
usufruct → UsUfruct	*subordinate*	*furrier*
cupric → cUpric	*uninformed*	*bulrush*

Another reliable rule pertains to certain vowels followed by a single consonant then a diphthong :h as *io*. As in rule (a), certain consonant-*r* combinations may play the role of a consonant.

$$[\underline{aeo}]\left\{\begin{array}{c}\neg[aehiouwxy]\\ th\end{array}\right\}\left\{\begin{array}{c}[ie]\\ ien\end{array}\right\}[\mathbf{aou}] \rightarrow [AEO]...$$

$$[\underline{aeo}][bcdfgkpt]ri[\mathbf{aou}] \rightarrow [AEO]...$$

Examples	Exceptions	Compare
sapient → sApient	*patio*	*recipient*
creation → creAtion	*special*	*division*
helium → hElium	*precious*	*beryllium*
cereal → cEreal	*national*	*forehead*
encomium → encOmium		*afield*
carpathian → carpAthian		*rapier*
specious → spEcious		*axial*
adriat|ic → Adriat|ic		*abreast*
appropriAte| → apprOpriAte|		*recreation*
tenebrious → tenEbrious		*alie|n*

A third reliable context for long vowels is *i* or *y* in a vowel pair before any other vowel in the word, it initial *y* is excepted:

$$\#\neg[\mathbf{aeiouy}]^*\left\{\begin{array}{c}\underline{i}\\ \neg[\mathbf{aeiouy}]\underline{y}\end{array}\right\}[\mathbf{aou}] \rightarrow ...\left\{\begin{array}{c}I\\ ...Y\end{array}\right\}...$$

Examples	Exceptions	Compare
iamb|ic → Iamb|ic	*kiosk*	*field*
biology → bIology	*kyoto*	*myel*
triUne| → trIUne|		*autobiography*
cyanIde| → cYanIde|		*symbiotic*
cryogen|ic → crYogen|ic		*yam*

ther contexts for long vowels are covered in the variable rules, for example *e* in *geology*.

5 Medial Silent E

Compound words whose first parts end with silent *e* cause special troubles. The following rules lo- ıte certain medial silent *e*'s with considerable reliability. They require that the *e* be isolated, and fol- ıwed by a syllable other than an ending that was marked during final *e* processing. Words ending in ɔnsonant-consonant-*le* can be recognized even when imbedded in compounds:

$$\neg[\#]\neg[\#][bcdfgmnprst][bdfgkpt]\underline{le}\neg[\mathbf{aeiouy}]\neg[|]^*[\mathbf{aeiouy}] \rightarrow ...le|...$$

lote: In applying this rule an appearance of *e*| is equivalent to | alone.

Examples	Exceptions	Compare
cockleshell → cockle|shell	*complementary*	*complEtion*
bumblebee → bumble|bee		*inflexible*
thistledown → thistle|down		*list|less|ness*

; is also often possible to recognize with fair confidence compounds whose first part is a canonical silent word.

$$\neg[aeiou\#][\underline{aiouy}]\neg[aeiouwxy]\underline{e}\neg[\mathbf{aeiouy}nr|]\neg[|]^*[\mathbf{aeiouy}] \rightarrow ...[AIOUY]...e|...$$

$$o[aiouy]\neg[aeiouwxy]\underline{e}\neg[\mathbf{aeiouy}nr|]\neg[|]^*[\mathbf{aeiouy}] \rightarrow ...e|...$$

'he same note applies as in (a).

Examples	Exceptions	Compare
racetrack → rAce|track	*malevolent*	*alphabet|ical*
houseboat → house|boat	*proletariat*	*paternity*
pacemAke|r → pAce|mAke|r		*relevant*
		inefficient
		alewife

These rules don't fully mirror the rules of Section 4.3 for final silent *e*. The differences are partly due to apparently skewed statistics of final *e* words that can be compounded, and partly to minimize the program for this relatively small class of words.

Though these rules do well as far as they go, they do not by any means recognize all embedded silent *e*'s. Here are some examples of undetected *e*'s.

edgeways	*horseback*
fencepost	*wherein*
largemouth	

4.6 Medial S

When it appears flanked by vowels, or preceded by a vowel and followed by *m, s* is usually voiced. Such an *s* is marked by capitalizing it:

(a) **[aeiouy]s[aeiouy]** → *...S...*

Examples	Exceptions	Compare
easy → eaSy	*andesite*	*pension*
desIre| → deSIre|	*osage*	*adversity*
invAsion → invASion		*version*
prismatic → priSmatic		*tUne|smith*

5 Tuning and evaluation

The rules arose from several sources—off the top of the head, the unabridged Webster[6], Venezky[5] and extensive interactive experimentation with several selected word lists. Certain difficult problems, such as the pronunciation of *ui* (compare *biscuit, ruin, anguish, guile, fruit, beguine*) were solved by perusing a 100 percent sample of the pertinent words in the unabridged Webster, which we had available on tape. Problems about patterns at the beginning of words were handled by consulting an ordinary dictionary[7]; patterns at the end of words were attacked with the help of a reversed-spelling dictionary[8]. These lists were used for tuning:

- Basic English[9]. This 900-plus word list is a rich source of common exceptions.
- A list of some 2800 most common words in a million-byte corpus of Bell Laboratories documents. The corpus was gathered and analyzed by R. Morris and L. L. Cherry.

The goal of tuning was to get pronunciations that are psychologically close to the range of American English heard day-to-day; there was no intent of approximating any one "standard" dialect. For example alternations between long and short version of *e* or of *o* are often acceptable; witness their occurrence in intelligible English spoken by foreigners. On the other hand such an alternation between long and short *u* baffles most listeners. Thus, we have little trouble understanding *defense* or *potential* even when their first syllables are pronounced as in *definition* or *potash*, while *munition* alliterated with *mundane* escapes us completely.

Once tuning was complete, the program's performance was evaluated against independent sources:

- The Brown Corpus[10]. The program was tested against the 2000 most common words, and an approximate one percent sample of the remainder to estimate the tail. With words frequencies taken into acount, performance was deemed satisfactory on about 97% of the 2000 most common words in running English, but only 88% of the tail, for an overall weighted performance of 97.2%.
- *Webster's Elementary Dictionary*[11]. A sample of one word per page, or about 3% of this 18,000-word source, had an error rate of about 5.5%.
- A 3552-word extract from Mark Twain[12]. This literary text has a substantially different vocabulary from the "Bell Labs English" on which the program was tuned, and is replete with startling references to distant subjects. As the listener, I was naive to the Twain, but thoroughly familiar with the idiosyncracies of the program. The sense of the passage was completely clear. I failed to recognize only 30 words, or about 0.8%. It turned out that many of these were pro-

nounced correctly within the limitations of the machine. Another 46 words received bizarre pronunciations, but were intelligible in context. Counted in approximately the same terms as was the Brown Corpus, the error rate on this document was about 2.1%.

To give an idea of the virtues and failings of the program, the list from the Elementary Dictionary appears in Appendix 4 partitioned into good pronunciations, clearly wrong, but intelligible pronunciations, and complete failures. The causes of failing are often transparent, but sometimes words go wrong as result of an attempt to get many others right. For instance *congest* was pronounced as if it were related to *longest*, *wronged*, etc., *screwdriver* was rhymed with *liver*, and *solemnity* was treated as a compound of *sole* and *mnity*. Some of the mispronunciations could be cleaned up at no cost; *heroic* could be fixed up by adding the rule *o*| → *O*, which would help a number of other words at the same time—for instance *go*|*ing*, which is in as special case. But you have to stop somewhere. Some chronic failings are harder to fix up; no less than 5 errors are attributable to embedded silent *e*'s of kinds the program can't recognize.

6 Discussion

6.1 Stress

Partly because the device at our disposal was deficient in this regard, but mostly because of the inherent difficulty of the problem, I have ignored stress in this work, except for three trivial experiments: (1)Placing falling inflection on syllables followed by periods. This fillip in itself does little for intelligibility. It is actually counterproductive in some cases, such as *Mr. Jones*, where it produces a major break within an intimately bound phrase. (2)Assigning one of four levels of stress to each syllable by a random walk algorithm. This has little effect on a wordwise basis, but definitely improves long passages by rendering them less deadening if not more natural. The effect is actually quite interesting when coupled with falling infection at sentence ends. (3)Giving correct stress to the known words, within the limitations of the machine. Sporadic correctly stressed polysyllables stand out against a background of monotone speech and actually detract from its apparent "ruliness", upon which one comes to depend. A systematic approach to the selection of words to be properly stressed, for example by choosing a comprehensive list of function words to be pronounced without stress, might help.

It would be possible to stress long vowels, for example the *a* in *oration* located by rule 4.4(b), but beyond this it would seem that considerable extension to the present program would be required to do a job of stress assignment even remotely approaching its performance on sounds. The incorporation of stress could even upset the compromises with pronunciation that turned out to be acceptable in unstressed speech. For example, stress would only call attention to the program's insignificant mispronunciation of *o* in *ovary* like *o* in *hot*. I suspect, too, that stress done badly is worse than no stress at all; but I could be quite wrong, for I have been surprised once already by how far you can get with pronunciation done wrong.

As in indicator of how much work might have to go into doing stress right, the stress assignment rules of Chomsky and Halle[13] appear to depend so intimately on an inferred context-sensitive grammar of English, as to frustrate any attempt to simulate them from simple analysis of spellings. Even the list of terminal symbols for that grammar is large—Allen[1] has made a lexicon of some 10,000 "morphs" covering only the Brown Corpus. I wouldn't be surprised if this lexicon had to be doubled to achieve good coverage of the Manhattan phone book, whose entries must count as good English words when the reading of running text is one's goal.

6.2 Improvability

The decision to take cognizance of no levels of English other than letters and words undoubtedly limits the growth potential of this approach to pronunciation. Its inattention to syntactic considerations precludes the program from doing any better than guessing at the pronunciation of words such as *read* or *bow*. Its disregard for morphology makes it difficult to do much better than presently on some kinds of compounds, such as *hothead*, though others can obviously be handled by adding morphs to the list of word fragments, as has already been done with *death*, *sOme* and many others. To overcome *hothead* difficulties, a certain rather small class of word fragments could be recognized fairly early on during preprocessing on other than a strictly left-to-right basis. More generally, to get much leverage out of further variable rules, it may be occasionally necessary to specify priorities among rules.

An idea of the expansibility of the present approach may be had from comparing it with Ainsworth's program for British English, which reportedly has about a 10% error rate[14]. In the terminology of this paper, that program appears to have the equivalent of about 3 preprocessing rules and 150 variable rules; the variable rules do not permit rewriting. The present program did about 4 times as well

on error rate at a price of about 4 times as much mechanism. To go further, often rules will have to be added to take care of just one or two words apiece. Much of the trouble will be with abbreviations and proper names, no one of which is very common, but the sum total of which will keep the error rate up against even considerable expansion of the rules.

6.3 Other Languages

Needless to say, the basic algorithmic methods of this paper work better on most other Western languages than they do on English. By changing the variable rules, and suppressing English preprocessing, we have obtained performance on Church Latin and on Polish comparable to that on English. In each case the whole tuning job took less than an hour. Unfortunately the commercial phoneme generator lacks come critical phonemes of most Western languages, most notably trilled *r*'s, gutturals and nasalized vowels, so it is infeasible at present to generate a comprehensive library of languages.

6.4 Utility

In small doses, the output of the program is usually intelligible to everyone who has heard very much of it. Short connected fragments out of context, sprung even when the hearer is not attending, are usually understood. Long passages, on the other hand, take extreme concentration to follow, and very short utterances of one or two syllables frequently suffer from the inherent limitations of our hardware. Even bizarre pronunciations are often decipherable, once the machine's "accent" has become familiar.

The program's outstanding virtue is its readiness to pronounce everything, albeit far-fetchedly on occasion. Of course its vocabulary can be hand-crafted to overcome any particular difficulty, but therein it becomes just a poor imitation of many other synthesis programs that do a far better job on hand-prepared input. Its speed and compactness make it attractive as an output routine for other programs, for example in making computers available to the blind (an application which should be distinguished from making computers substitute for talking books).

My basic reason for setting forth the details of the program here, however, is not to advertise the particular implementation, but rather the belief that the approach may be easily copied as the method of last resort in other synthesis projects that depend on dictionaries of words, stems, affixes or whatever. An authentically comprehensive dictionary for arbitrary text is not yet in sight physically, say nothing of economically for many applications. For this purpose, the rules of this paper may go a bit too far towards handling particular exceptional words, but it is always easier to take rules away than to put them in. With suitable, but usually obvious, modifications of the phonetics, the rules should be quite widely applicable.

Speculating further, one wonders whether this basically "stupid" approach to synthesis is viable in the long run against linguistically more sophisticated attacks. To what extent does it mirror a child's ability to read material containing words he can't understand? How is a dictionary impaired if obviously pronounceable words are omitted? As a matter of economy, how much word-level grammatical sophistication should be in a speaking program? Is it cheaper to cobble more "intelligence" into a babbling idiot, or to stuff a vocabulary into a dumb grammarian? At this juncture, I don't pretend to know the answers; but I don't believe anybody else knows either.

7 Acknowledgements

Several people devoted many hours to making the voice synthesizer work in our laboratory, and learning how to play it. K. Thompson and D. M. Ritchie integrated the device smoothly into the operating system. J. F. Ossanna was the first to make it say things, and also saw the hardware through a series of debilitating diseases. R. Morris and L. E. McMahon spent several hundred man-hours developing phonetic codings and hand tuning an initial vocabulary approaching 1000 words. Without the benefit of their experience, and frequent application of their perceptive ears, the tuning of the final program would have suffered immensely. Discussions with my first-grade son, Peter, who was intensely sympathetic to the difficulties of pronouncing English you don't fully understand, were also illuminating.

A number of UNIX software tools proved invaluable to the development work, most notably the pattern-selecting program 'grep', which K. Thompson wrote especially to help with dictionary searches. The machine-readable dictionary and word lists from which pronunciation rules were derived, and with which verification and testing were done, were put together by M. R. Wagner, D. M. Ritchie, R. Morris, L. L. Cherry and L. E. McMahon. This paper itself would probably have fallen between two stools, proofread but unprintable, or printed and unproofread, had it not been possible to compare the rules and examples of the Appendices mechanically against the program; J. F. Ossanna's 'troff' made it possible to solve the special style problems of the paper, while keeping it machine-readable.

Bibliography

[1] Votrax, Vocal Interface Division, Federal Screw Works

[2] C. H. Coker, N. Umeda and C. P. Browman, Automatic Synthesis from Ordinary EnglishText, *IEEE Trans. Audio and Electroacoustics*, AU-21(1973) 293-298

[3] J. Allen, Reading Machines for the Blind: The Technical Problems and the Methods Adopted for Their Solution, *IEEE Trans. Audio and Electroacoustics*, AU-21(1973) 259-264

[4] D. M. Ritchie and K. Thompson, The UNIX Time-Sharing System, *CACM* to appear

[5] R. L. Venezky, *The Structure of English Orthography*, Mouton, Paris 1970

[6] *Webster's New International Dictionary*, 2nd edition, G. & C. Merriam, Springfield 1958

[7] *Webster's Seventh New Collegiate Dictionary*, G. & C. Merriam, Springfield 1970

[8] L. H. Dawson, *Walker's Rhyming Dictionary*, Dutton, New York 1936

[9] C. K. Ogden, *Basic English*, London 1940

[10] H. Kucera and W. N. Francis, *Computational Analysis of Present-Day American English*, Brown University Press, Providence 1967

[11] *Webster's Elementary Dictionary*, G. & C. Merriam, Springfield 1966

[12] S. L. Clemens, Ascending the Riffelberg, in *A Tramp Abroad*

[13] N. Chomsky and M. Halle, *The Sound Pattern of English*, Harper and Row, New York 1968

[14] W. A. Ainsworth, A System for Converting English Text into Speech, *IEEE Trans. Audio and Electroacoustics*, AU-21(1973) 288-290

Appendix 1. Complete Words

These are the only words tabled *qua* words. All other words are pronounced by rule. No phonetics are given here, for it seems inappropriate to burden the paper with the exact phonetic transcription for one unique device.

Many of these words are notorious exceptions in English. Others are more subtle: *as, gas* and *has* to prevent them from rhyming with *was, over* with *cover, any* and *many* with *zany, shall* with *hall, water* with *later.* Some words are on the list because the rules simply run amuck on them—*something* would otherwise become *somETH|ing, yes* would become *Ies.*

a	*doing*	*has*	*mr*	*seven*	*tuesday*
alkali	*done*	*have*	*mrs*	*shall*	*two*
always	*dr*	*having*	*nature*	*someone*	*upon*
any	*early*	*heard*	*none*	*something*	*very*
april	*earn*	*his*	*nothing*	*than*	*water*
are	*eleven*	*imply*	*nowhere*	*that*	*wednesday*
as	*enable*	*into*	*nuisance*	*the*	*were*
because	*engine*	*is*	*of*	*their*	*who*
been	*etc*	*island*	*on*	*them*	*whom*
being	*evening*	*john*	*once*	*there*	*whose*
below	*every*	*july*	*one*	*thereby*	*woman*
body	*everyone*	*live*	*only*	*these*	*women*
both	*eye*	*lived*	*over*	*they*	*yes*
busy	*february*	*living*	*people*	*this*	
copy	*finally*	*many*	*read*	*those*	
do	*friday*	*maybe*	*reader*	*to*	
does	*gas*	*meant*	*refer*	*today*	
doesn't	*guest*	*moreover*	*says*	*tomorrow*	

Appendix 2. Rules for Word Fragments

The fragments are written as they occur after preprocessing, or as a result of rewriting other fragments. The examples are written as after preprocessing. The mark † denotes a poor but intelligible pronunciation. The comparison words are intended to suggest why plausible alternative rules might not work, evidently related cases that a given rule misses, putative exceptions that are forestalled by preprocessing, or related rules which are good candidates for adding to the list. The rules are arranged hierarchically, where deeper levels express exceptions to shallower levels. Thus rules 1, 1.5 and 1.5.1 tell that *A* is usually pronounced long, but not in words like *rATHe|r*, among which class *fATHe|r* needs special treatment.

	Rule		Examples	Exceptions	Compare
1	*A*	*ai,y2*	*mAne| capAcious main*	*glissAde| immEdiAte|† morAle| pavAne| senAte| thAme|s ultrAviolet*	
1.1	*A#*	*e1,y1*	*play* 2.16		
1.2	*Ar*	*ai→r*	*bAre|ly air*		
1.3	*Ad|or*	*u2→dor*	*ambassAd|or*		*monIt|or* 9.1 *rAdiAt|or*
1.4	*Anion*	*→anion*	*compAnion*	*Anion*	
1.5	*ATHe|r*	*→aTHer*	*gATHe|r rATHe|r*	*bATHe|r*	*whETHe|r* 6.3
1.5.1	*fATHe|r*	*f,a0→THer*	*grandfATHe|r*		*fatherhood*
1.6	*{nr}Ational*	*→{nr}ational*	*nAtionaliSm rAtional|ly*	*vibrAtional*	*nAtion orAtion*
1.7	*#privAte*	*→prIvite*	*privAte|ly*		*deprivAtion privatee|r privAtion*
2	*a*	*ae*	*cat village|*	*already† oracle|† scarce| saar vacAte|*	
2.1	*a#*	*u2*	*algebra india*		
2.2.1	*|abl*	*u2,b,l*	*const|able palp|ably*	*dIs|able unsr|able*	*tAble*
2.2.2	*able|*	*→Able|*	*able|r enable|d*		
2.3	*ae*	*e1*	*aero- aesthet|ic*	*gael|ic paean*	*encyclopaEdia*
2.3.1	*ae|#*	*→E*	*alumnae| fringillidae|*		
2.3.2	*aeo*	*→Eo*	*palaeolith|ic aeon*		
2.4	*#agen*	*→Agen*	*agency agent*	*agenda*	*magenta*
2.5	*ague*	*ai,y1,g→e*	*vague|*	*daguerrotYpe| montague|*	
2.6	*ah#*	*a0*	*hurrah shah*	*fellah sarah*	*ahead autobahn*

	Rule		Examples	Exceptions	Compare
2.7	*ai*	*→A*	*again* *air* *pertain* *raiSe|*	*aisle|* *bargain* *plaid*	
2.7.1	*a|ic*	*→Aic*	*hebra|ic*		
2.7.2	*certain*	*→serte|n*	*uncertain* *certainty*	*ascertain*	*curtain* *mountain* *villain*
2.7.3	*said*	*→sed*	*afOre|said*		
2.8.1	*al[#|]*	*u2→l*	*trIal* *eventuat|ly*	*ital|ic* *tal|ly*	
2.8.2	*ald*	*aw,l,d*	*bald* *alderne|y*	*aldehYde|* *herald†*	
2.8.3	*alf#*	*ae,f*	*half*		*alfalfa*
2.8.4	*alk*	*→awk*	*talk*	*balkan†*	
2.8.4.1	*#alk*	*ae,l,k*	*alkalIne|*		
2.8.5.1	*all#*	*→awl*	*ball*		*hallow* *alliance|*
2.8.5.2	*alle|*	*→awle|*	*falle|n* *smalle|r*	*galle|ry*	
2.8.5.3	*all|ing*	*→awling*	*call|ing*		*gall|ic*
2.8.5.4	*alls#*	*aw→ls#*	*falls*		
2.8.6	*alm*	*a0→m*	*calm* *almond*	*almanac* *salmon†*	
2.8.7	*al[st]*	*aw,l→[st]*	*alternAte|* *also* *false|* *salt*	*alto*	
2.9	*#any*	*e0,n,ie,y1*	*anyhow*		*Any*
2.10	*ar*	*a0,er*	*artist* *far*	*parOle|†* *singUlar†*	
2.10.1	*ar[aeio]*	*ea→r[aeio]*	*paradISe|* *parent* *marinAte|* *faro*	*caress†*	*america* 7.12.2 *miracle|* 10.11.1
2.10.1.1	*arou*	*u2→rou*	*around* *carouSe|*	*viviparous*	
2.10.2	*arr*	*ae→r*	*marry*	*dIarrhea†* *quarry*	*arrow* 17.16.5.2.1
2.10.3	*ary*	*ai,r,y1*	*binary* *stAtionary*		
2.11	*ao*	*→Ao*	*aorta* *kaolin*	*gaol*	
2.11.1	*ao#*	*→ow*	*mao*		
2.12	*as#*	*u2→s#*	*americas*	*atlas†* *canvas†*	
2.13	*aste|*	*→Aste|*	*haste|n* *waste|*	*aste|r*	*hasty* *nasty* *wast|ing*
2.13.1	*[cflm]ast*	*[kflm],ae→st*	*caste|* *faste|n* *blaste|d* *maste|r*		

	Rule		Examples	Exceptions	Compare
2.14	*au*	*→aw*	*auTH|or* *haul* *taught*	*aunt* *chauffeur* *gauge|* *sauerkraut*	
2.14.1	*#laugh*	*→laf*	*laughte|r*		*slaughte|r*
2.15	*aw*	*aw*	*awkward* *drawn*	*Areawlde|*	
2.15.1	*aw#*	*u2,aw*	*draw*		
2.15.2	*aw[Aa]*	*u2→w[Aa]*	*UnawAke|* *award*		
2.16	*ay*	*→A*	*bay* *mayhem*	*aye|* *kayak*	
2.17	*adjac*	*u2→jAc*	*adjacency*		
2.18	*bas|ic*	*→bAsic*	*bas|ical|ly*		
2.19.1	*change*	*→chAnje*	*change|d*		
2.19.2	*chang|*	*→chAnj*	*chang|ing*		
2.19.3	*[dr]ange*	*→[dr]Anje*	*dange|r* *range|d*		
2.19.4	*rang|i*	*→rAnji*	*derang|ing*		
2.20	*label*	*→lAbl*	*label|ing*		
2.21.1	*place|*	*→plAce|*	*displace|*		*complacent* *placenta*
2.21.2	*plac|i*	*→plAci*	*replac|ing*		*placid*
2.22	*wa*	*w→aw*	*awash* *want*	*swallow†* *swat*	
2.22.1	*wa[gixy]*	*w→a[gixy]*	*swagge|r* *wait* *wax* *way*		
2.22.2	*walk*	*w→alk*	*norwalk*		
2.22.3	*wang*	*w→ang*	*twang* *wangle|*		
2.22.4	*war*	*w,o1,r*	*quarrel* 19.1 2.22 *forward†* *swarthy* *warllke|*		*howard* 17.16
2.23	*what*	*h,w,u1,t*	*whatEve|r*		
3	*b*	*b*	*baobab* *lumbe|r*	*bombe|r* *bomber* *dumb†* *subpoena*	
3.1	*bb*	*→b*	*bubble|*		
3.2	*bt*	*t*	*debt* *doubt* *subtle|*	*subtend* *bobtail*	
3.2.1	*obt*	*a0,b→t*	*obtain* *obtUSe|*		
4	*c*	*k*	*cataclySm|ic* *success*	*caeSar* *czar* *czech* *facAde|* *indict*	

	Rule		Examples	Exceptions	Compare
4.1	*c[EeIiY]*	*→s[EeIiY]*	*precEde| success placid incIte| cYcle*		*cello receive|* 7.8.4
4.1.1	*ce[ao]*	*→ci[ao]*	*Ocean herbAceous*		
4.1.2	*c|i*	*→si*	*fenc|ing*		
4.1.3	*ci[aeo]*	*sh,ih→[aeo]*	*fAcial efficient vicious*		*conscience|* 22.4
4.2	*ch*	*t,ch*	*church*	*ache| cache|† chaSm chef choir chthOnian drachm echo fuchsia parachUte| stomach yacht*	
4.2.1	*charact*	*k,ae,r,i2,k,t,t*	*characterIze|*		*charAde| charm*
4.2.2	*chem*	*→kem*	*chem|ical*	*sachem chemISe|*	
4.2.3	*.zn 4.2.4 ch[lnr]*	*k→[lnr]*	*chlorIde| technique| achromat|ic*		*match|less such|ness*
4.2.5.1	*chor*	*→kor*	*chordAte|*	*chortle|*	*chOre| chOrus*
4.2.5.2	*ch|or*	*→k|or*	*anch|ore|d parach|or*		
4.2.6	*arch[iy]*	*→ark[iy]*	*architect hierarchy*		*archIve| pAtriarch archdUke|*
4.2.7	*machIn*	*m,a1,sh→En*	*machIne|ry*		*machinAtion*
4.2.8	*mech*	*→mek*	*mechaniSm*		
4.3	*ck*	*k*	*check|ing acknowledge|*		
4.4	*cy*	*s,s→y*	*decency*		*cyclame|n* 30.4 *pansy* 22.9
5	*d*	*d*	*indee|d*	*soldie|r individual†*	
5.1	*d#*	*d,t*	*hard wishe|d*		
5.2	*dd*	*→d*	*fiddle| padd|ing*		
6	*E*	*ie,y0*	*concrEte| sincEre|*	*fEte|*	
6.1	*E#*	*y0,y1*	*fee|*		

	Rule		Examples	Exceptions	Compare
6.2	*E[dt]al*	*→e[dt]al*	*pEdal* *mEtal*	*fEtal*	
6.3	*ETHe|r*	*→eTHer*	*whETHe|r*	*ETHe|r*	*rATHe|r* 1.5 *leaTHe|r* 7.4.3 *hITHe|r* 9.2 *OTHe|r* 16.1 *ethEreal*
6.4	*Eve|r*	*e1,v,er*	*nEve|r* *sEve|re|d*	*fEve|r* *rEve|re|d*	*feverish* *reverent*
6.5	*amEte|r*	*→ameter*	*diamEte|r* *paramEte|r*	*decamEte|r*	
6.6	*discrEtion*	*→discretion*	*discrEtionary*		*concrEtion* *excrEtion*
6.7	*prEcious*	*→precious*			
6.8	*spEcial*	*→special*	*espEcial|ly*		
7	*e*	*e0*	*best* *impel* *peddlar*	*beret* *england* *hedgehog* *renege|* *travel†*	*begin* 8.2.2.6
7.1	*e#*	*→*	*Acre*		
7.2	*e|*	*→*	*mAde|* *wishe|d*	*psyche|*	*grandee|*
7.3	*é*	*→A*	*glacé* *née|*		
7.4	*ea*	*ie,y0*	*breaTHe|* *congeal* *fear* *sheaf* *teak* *weary* *wreath*	*beatrice|* *breath* *deaf* *linear* *rOSeate|* *steak* *vengeance|* *wear*	*idea* 10.4
7.4.1.1	*earch*	*er,ch*	*rESearch*		
7.4.1.2	*ear[ln]*	*→er[ln]*	*pearl* *learn*		*near|ness*
7.4.1.3	*ea[rl]th*	*→e[rl]th*	*Unearthe|d* *wealthy*		
7.4.1.4	*heard*	*→herd*	*unheard*		
7.4.1.5	*heart*	*→hart*	*hearth* *hearty*		
7.4.2	*eaSUr.*	*e1,zh→Ur*	*meaSUre|ment* *treaSUry*		
7.4.3	*eaTHe|r*	*→eTHe|r*	*leaTHe|ry* *weaTHe|r*	*breaTHe|r*	*weatherman* *whETHe|r* 6.3
7.4.4	*eau*	*→O*	*beau* *bUreau*		
7.4.4.1	*beaut*	*→bUt*	*beauti|ful*		
7.4.5	*Area*	*ai,r→Ea*			
7.4.6	*bear*	*→bAr*	*unbear|able*	*beard*	
7.4.7	*break*	*→brAk*	*unbreak|able*	*breakfast*	

	Rule		Examples	Exceptions	Compare
7.4.8	*creat*	*→crEAt*	*create|* *creat|ing*	*creatUre|*	
7.4.9	*dead*	*→ded*	*dead|ly*		
7.4.10	*death*	*→deth*	*death|less*		
7.4.11	*great*	*→grAt*	*great|ly*		
7.4.12	*head*	*→hed*	*headlong* *fOre|head*		*hOthead*
7.4.13	*heav*	*h,e1,v*	*heave|n* *heavy* *heavie|r*	*heave|*	*sheave|*
7.4.14	*react*	*→rEact*			
7.4.15	*read*	*r,e1,d*	*bread* *readi|ly*	*reads*	
7.4.15.1	*#read|*	*→rEd*	*read|able* *read|ing*		
7.4.16	*rhea*	*→rEa*	*logorrhea*		
7.4.17	*stead*	*→sted*	*instead* *steady*		
7.4.18	*sweat*	*→swet*	*sweate|r*		
7.4.19	*threat*	*→thret*	*threate|n*		
7.5	*ecent*	*→Ecent*	*decent* *recent|ly*		
7.6	*e|ct*	*→ect*	*dIre|ction*		
7.6.1	*e'd*	*→Ed*	*he'd* *she'd* *we'd*		
7.6.2	*[dt]e|d*	*[dt],i2,d*	*miste|d* *winde|d*	*miste|d*	*crooke|d* *indee|d*
7.6.2.1	*e|d!ly*	*i2→dty*	*decIde|d|ly* *heate|d|ly*		
7.7	*ee*	*→E*	*pee|r* *feeble|* *nominee|*	*breeche|s* *matinee|* *reelect*	
7.8	*ei*	*→A*	*heir* *rein* *seine|* *weight*	*deity* *eide|r* *heife|r* *leiSUre|* *protein* *seize|* *sleight* *villein* *weird*	
7.8.1	*eign*	*→An*	*reign*	*sovereign*	
7.8.1.1	*foreign*	*f,o0,r,e2,n*	*foreigne|r*		
7.8.2	*eiTHe|r*	*→ETHer*	*neiTHe|r*		
7.8.3	*beit#*	*→bEit#*	*albeit* *howbeit*		
7.8.4	*cei*	*→sE*	*receive|*		
7.8.5	*feit*	*f,i2→t*	*forfeite|d* *surfeit*		

	Rule		Examples	Exceptions	Compare
7.8.6	*heigh*	*→hI*	*height*		
7.8.7	*stein*	*→stIn*	*bernstein* *steinberg*		*einstein* *heiSenberg*
7.9	*e|n*	*e2,n*	*rAve|n* *heave|n|ly*		*see|n* 7.7
7.10	*eo*	*→Eo*	*geology* *rheostat*	*george|* *leopard* *pigeon* *yeoman*	*gorgeous* 8.2.2.1
7.11	*#eq*	*→Eq*	*equal* *equestrian*		
7.12	*er*	*u2,er*	*her* *ferOcious* *perhaps*	*concerto* *herewith* *sergeant* *rerun* *zero*	*therein* 23.5.1
7.12.1	*e|r*	*er*	*dEte|r* *highe|r*		
7.12.2	*eri*	*e1,r→i*	*america*		
7.12.3.1	*er|i*	*→eri*	*homer|ic*		*hamme|r|ing*
7.12.4	*err*	*e1,r*	*berry* *errand* *terrible|*		
7.12.4.1	*erre|*	*er→e|*	*inferre|d*		
7.12.5	*wh[Ee]re*	*→whAre*	*sOme|whEre|* *wherewithal*		
7.13.1	*es#*	*→s#*	*Acres*		
7.13.2	*e|ss*	*→ess*	*prOce|ssion*		
7.13.3	*e|s[qt]*	*e1→s[qt]*	*dIve|stitUre|* *fIne|st* *pictUre|sque|*	*fOre|stay*	*reque|st* 19.1.1.1 *west*
7.13.4.1	*che|s#*	*ch,i2→s#*	*inche|s*		
7.13.4.2	*je|s#*	*d,j,i2→s#*	*flange|s* *rAge|s*		
7.13.4.3	*se|s#*	*s,s,i2→s#*	*bASe|s* *rAce|s*		
7.13.4.4	*she|s#*	*sh,i2→s#*	*fishe|s*		
7.13.4.5	*ze|s#*	*z,i2→s#*	*buzze|s*		
7.13.5	*phe|s#*	*→fEs#*	*catastrophe|s*		
7.14	*eu*	*→ew*	*eugen|ics* *neute|r*	*sabOteur*	*lieu* 10.5.5
7.15	*ew*	*yI,yu*	*few* *hewn* *pewte|r*	*bewAre|†*	
7.15.1	*[djlrst]ew*	*→[djlrst]U*	*dew* *jewel* *flew* *grew* *sewe|r* *stew*	*alewIfe|* *rewind* *sew*	*dUty* 24.4
7.16	*ey*	*→A*	*abeyance|* *grey*	*boneyard* *geySe|r* *honeycomb* *jerseyIte|*	

	Rule		Examples	Exceptions	Compare
7.16.1	*e|y*	*→E*	*chimne|y* *jerse|y* *monke|y*	*Obe|y*	
7.16.1.1	*ve|y*	*→vey*	*conve|y* *surve|y*	*cOve|y* *slAve|y*	
7.16.2	*key*	*→kE*	*key* *keyboard*		
7.17	*phe|#*	*→fy#*	*apostrophe|*		*gophe|r* *strophe|s* 7.13.5
7.18	*ple|ment*	*p,l,u2→ment*	*comple|ment* *imple|ment*		
7.19	*#prett*	*→prit*	*pretty*		
7.20	*prOce|d*	*→prOcEd*	*prOce|dUral*		*prOce|ssion*
7.21	*secret*	*→sEcret*	*secret|ly*		
8	*f*	*f*	*fifth*		
8.1	*f#*	*f,f*	*chief* *staff* *tough*		
8.1.1	*ff*	*→f*	*raffle|* *staff*		
8.2	*g*	*g*	*enigma* *gag*	*algae|* *champagne|* *margarIne|* *paradigm*	
8.2.1	*gg*	*g*	*digge|r* *haggle|*		
8.2.1.1	*exagg*	*→exag*	*exaggerAte|*		
8.2.1.2	*sugg*	*→sug*	*suggestion*		
8.2.2	*g[EeIiy]*	*→j[EeIiy]*	*gEnius* *gIant* *rEgion* *edgy*	*bOgy* *eage|r* *garage|†* *gear* *gecko* *geeSe|* *gift* *gill* *tIge|r*	*binge|* 15.3.1.1 *finge|r* 15.3.1.1.1
8.2.2.1	*geous*	*d,i,ih→ous*	*gorgeous*		
8.2.2.2	*forget*	*→forgget*	*forgett|ing*		*energet|ic* *gorget* *target*
8.2.2.3	*#get*	*g,e1,t*	*getaway*		*midget* *vegetAte|*
8.2.2.4	*girl*	*g,er,l*	*girlish*		
8.2.2.5	*gIv*	*g,i1,v*	*gIve|n* *thanksgIv|ing*	*ogIve|*	
8.2.2.6	*begin*	*→bEggin*	*beginn|ing*		
8.2.3	*gh*	*→*	*taught* *weigh*	*aghast* *dinghy* *newburgh* *spaghetti*	*laugh* 2.14.1 *rough* 17.14.3 *high* 10.7

	Rule		Examples	Exceptions	Compare
8.2.3.1	*#gh*	*→g*	*ghetto* *ghast*		
8.2.4	*g|i*	*→ji*	*mag|ic* *rAg|ing*		*bang|ing* 15.3
8.2.5	*glio*	*→lio*	*imbroglio*		
8.2.6	*#gn*	*n*	*gnat* *gnU*		*agnew* *arraign*
8.2.6.1	*ign*	*→In*	*malign* *reSign|ing*		
8.2.6.1.1	*ign\[AaeIiOo\]*	*i0.g→n\[AaeIiOo\]*	*indignAtion* *malignant* *igneous* *ignIte|* *dignity* *ignOre|* *ignoble|*	*mignon* *signout* *vignette|*	
8.2.6.1.1.1	*igne|*	*→Ine|*	*reSigne|d*		
8.2.7	*h*	*h*	*hitchhIke|*	*dinghy* *haverhill* *john* *khan* *hour†* *hOne|st* *wAre|ham*	*ah* 2.6
8.2.7.1	*hOn|or*	*→on|or*	*hOn|or|able*		*honorary*
9	*I*	*a0,iy*	*hIre|* *valentIne|*	*dIs|able* *medicIne|* *fertIle|†* *jUpIte|r* *ravIne|* *regIme|* *skIs* *vallSe|*	*machIne|* 4.2.7
9.1	*I\[dt\]|or*	*→i\[dt\]|or*	*corrId|or* *monIt|or*		*ambassAd|or* 1.3
9.2	*ITHe|r*	*→iTHer*	*whITHe|r*		*whETHe|r* 6.3
9.3	*Ity*	*→ity*	*cIty* *pIty*		
9.4	*Ive|r*	*→iver*	*IIve|r* *rIve|r*	*drIve|r* *fIve|r*	*drIve|n* *IIve|n*
9.5.1	*SI\[tv\]*	*→si\[tv\]*	*depoSIte|d* *oppoSIte|* *intrUSIve|*	*andeSIte|* *paraSIte|* *viSIt|or†*	*sIte|*
9.5.2	*\[st\]Iv*	*→\[st\]iv*	*intensIve|* *preventIve|*		*gIve|* 8.2.2.5
9.5.3	*\[aeio\]SIv*	*→\[AEIO\]siv*	*evaSIve|* *adheSIve|* *diviSIve|* *exploSIve|*		
9.6	*consIde|r*	*→consider*	*consIde|r|able*		*insIde|r*
9.7	*determIn*	*→dEtermin*	*undetermIne|d*		
9.8	*edIt*	*→edit*	*edIt|ing*	*expedIte|*	

	Rule		Examples	Exceptions	Compare
9.9	*#examIn*	*→examin*	*examIne|d*		
9.10	*finIt*	*→finit*	*definIte|ly*		
9.10.1	*#finIt*	*→fInIt*	*finIte|ly*		
9.11	*limIt*	*→limit*	*unlimIte|d*		
9.12	*routIn*	*→rUtEn*	*routIne|*		
10	*i*	*i0*	*austria* *infinitUde|* *similar* *vAriant* *vestige|*	*alliance|* *climb* *disciple|* *familiar†* *graffiti* *marriage|†* *meringue|* *parlia|ment* *prestige|* *silence|*	*sign* 8.2.6.1 *deriSIve|* 9.5.3 *rAtio* 23.6.4
10.1	*i#*	*y1*	*ennui* *stokowski*	*alibi*	
10.2	*#i'*	*→1*	*i'd* *i'll*		
10.3.1	*ible|#*	*i2,b,l*	*terrible|*		*bIble* *sible|y*
10.3.2	*ibly*	*i2,b,l,y1*	*sensib|ly*		
10.4	*idea*	*→IdEa*	*idealiSm*		
10.5	*i[Ee]*	*ie,y0*	*reliEve|* *carrie|d* *field*	*diEne|* *fiesta* *siEve|*	
10.5.1	*ien*	*y0→en*	*convEnient*	*fiend*	*alie|n* *IIe|n*
10.5.1.1	*friend*	*→frend*	*unfriend|ly*		
10.5.2	*ie|r*	*iy,y1,er*	*IOne|tie|r*	*collie|r†* *dossie|r*	*fierce|*
10.5.3	*ie|st*	*y1,u2,s,t*	*dirtie|st*		*prIe|st*
10.5.4	*iet*	*→Iet*	*diet* *quiet*	*soviet*	
10.5.4.1	*ieth*	*→Eeth*	*fortieth*	*lieth*	
10.5.5	*ieu*	*iy,iu,iu*	*lieu* *milieu*		
10.5.6	*hier*	*→hIer*	*hierarchy*		
10.5.7	*ifie*	*→ifIe*	*amplifie|d* *petrifie|d*		*amplify* 30.6
10.5.8	*isfie*	*→isfIe*	*satisfie|d*		*satisfy* 30.7
10.5.9	*plie|*	*→plIe|*	*implie|d* *supplie|r*		
10.6	*i[gq]ue|*	*→E[gk]|e|*	*intrigue|* *Unique|*		*vague|* 2.5 *morgue|* 25.4.1.1 *intrigu|ing* *deliquescent*
10.7	*igh*	*→1*	*high* *mighty*		*eight* 7.8
10.8	*ild#*	*→IId*	*child* *mild*		*childre|n* *milde|r* *mildew* *wilde|r|ness*

	Rule		Examples	Exceptions	Compare
10.9	*ind*	*→Ind*	*blind|ly* *mind*	*kindre|d* *windlass*	
10.9.1	*ind\[eiou]*	*→innd\[eiou]*	*tinde|r* *indigent* *indoctrinAte|* *industry*	*finde|r*	*indOle|* *indUbIt|able*
10.9.2	*indl*	*→inndl*	*swindle|r*		*mind|less*
10.10	*io*	*y2→o*	*carrion* *dUbious*	*nonviolent*	*fAcial* 4.1.3 *lIon* *passion* 22.8
10.10.1	*io#*	*y1→o#*	*stUdio*	*ohio*	
10.11	*ir*	*u2,r*	*fir* *thirsty*		
10.11.1	*ir\[ai]*	*y1→r\[ai]*	*miracle|* *syrian* 30		*pirogue|*
10.11.2	*ir|i*	*y1→ri*	*empir|ical*		*er|ic* 7.12.3.1
10.11.3	*irr*	*y1,r*	*irreverent*		
10.11.4	*iron*	*a0,iy,er,n*	*flatiron*	*adirondack* *iron|ic*	
10.12	*#item*	*a0,iy,t,e1,m*	*itemIze|*		
10.13	*itis#*	*→Itis*	*tonsillitis*		*british*
10.14	*iu*	*y1→u*	*hElium* *rAdius*		
10.15	*#libr*	*→lIbr*	*librAtion* *library*	*librium*	*calibrAtion* *vibrAtion*
10.16	*#micr*	*→mIcr*	*microscOpe|*		*mimicry*
10.17	*satisfi*	*→satisfI*	*satisfie|d*		*misfit*
10.18	*titl*	*→tItl*	*entitle|*		
11	*j*	*d,zh*	*judge|*	*jaege|r* *juan* *rijkstafel*	
12	*k*	*k*	*kUlak*		
12.1	*kn*	*→n*	*knee|*		
13	*l*	*l*	*loyal* *salve|r*	*colonel* *lincoln* *salve|*	*calm* 2.8.6 *half* 2.8.3 *talk* 2.8.4 *would* 17.14.4
13.1	*ll*	*→l*	*million* *villa*	*llama*	
13.2	*l|l*	*→l*	*eventual|ly*		
14	*m*	*m*	*madam*	*comptrolle|r* *mnemon|ic*†	
14.1	*mm*	*→m*	*hamm|ing*		
15	*n*	*n*	*nancy*	*autumn*† *kiln*†	
15.1	*nn*	*n*	*mann|ing*		
15.2	*n\[ckx]*	*ng→\[ckx]*	*punctUre|* *sink* *lynx*	*unknown*†	
15.2.1	*nc\[eiy|]*	*n→c\[eiy|]*	*fence|* *invincible|* *nancy* *danc|ing*		

	Rule		Examples	Exceptions	Compare
15.3	*ng*	*ng*	*bang* *songfe|st* *stringy*	*stingy*	*song* 17.8
15.3.1	*ng{aEeiy}*	*n→g{aEeiy}*	*engage|* *ingEnious* *congenital* *enginee|r* *mangy*	*thingamajig* *tangy*	
15.3.1.1	*nge|*	*n→ge|*	*hinge|* *plunge|s*		
15.3.1.1.1	*nge|{dr}*	*ng,g→e|{dr}*	*bange|d* *finge|r* *longe|r*	*ginge|r* *mange|r* *singe|d*	*dange|r* 2.19.3
15.3.1.1.2	*nge|st*	*ng,g→e|st*	*longe|st*	*inge|st*	
15.3.2	*ng{loU}*	*ng→g{loU}*	*mangle|* *tango* *singUlar*	*ungod|ly†*	
15.3.3	*ngu*	*ng,g,w*	*distinguish* *language|*	*engulf* *tongue|*	
16	*O*	*o0*	*hOle|* *mOtion*	*bigOte|d* *dOze|n* *lOSe|* *Ove|n* *Onion*	*pensiOne|r* 22.8
16.1	*OTHe|r*	*→oTHer*	*brOTHe|r* *OTHe|r*	*bOTHe|r*	
16.2.1	*Ove|r*	*u0,v,er*	*cOve|r* *hOve|r*	*clOve|r*	*govern|ment* *overthrow* 17.15 *Ove|n*
16.2.2	*abOve*	*u2,b,u0,v*	*abOve|board*		
16.2.3	*lOv*	*→luv*	*glOve|* *lOve|ly*	*clOve|*	*shOve|*
16.2.4	*mOv*	*m,iu,iu,v*	*mOve|ment* *remOv|able*		
16.2.5	*prOv*	*p,r,iu,iu,v*	*improOve|ment*		*improviSAtion*
16.3	*cOl|or*	*→cul|or*	*discOl|ore|d*		*colorAtion*
16.4	*cOm*	*→kum*	*cOm|ing* *incOme|*		*comedy* *sOme|* 16.10
16.5	*gOne*	*→gawne*	*fOre|gOne|* *gOne|r*	*antigOne|* *orgOne|*	
16.6	*mmOn*	*m,u2,n*	*commOne|r*		
16.7	*mOne|y*	*→mune|y*	*mOne|y*		*hOne|y* *monie|d*
16.8	*prOpe|r*	*→proper*	*improOpe|r|ly*		
16.9	*purpOSe*	*p,er,p,u2→se*	*purpOSe|ful* *purpOSe|s*		
16.10	*sOme*	*→sume*	*lOne|sOme|* *sOme|whEre|*		*cOme|* 16.4 *riboSOme|* *somerset*
16.11	*SOn*	*→Son*	*reaSOn|able* *poiSOne|d*	*capariSOne|d†* *cortiSOne|*	*jacksOnian*

	Rule		Examples	Exceptions	Compare
16.12	*velOp*	*v,e0,l,u2,p*	*develOp|ing* *envelOpe|d*	*envelOpe|*	*develop|ment*
17	*o*	*a0*	*concoct* *catalogue|* *golf*	*[bpt]oll* *boSom* *cloSUre|* *cost†* *front†* *oh* *rogue|* *tomb* *wolf*	*ploSIve|* 9.5.3
17.1	*o#*	*o0,o1*	*largo*		
17.2	*oa*	*o0*	*cocoa* *oat*	*protozoa* *oaSis*	
17.2.1	*broad*	*b,r,aw,d*	*abroad*		
17.3	*oe*	*→Oe*	*coerce|* *hoe|* *poem*	*amoeba* *coeval*	*doe|r* *hoedown*
17.3.1	*shoe*	*→shooe*	*shoehorn*		*hoopoe|*
17.4	*of[ft]*	*aw→f[ft]*	*coffee|* *soft*		
17.4.1	*ofte|n*	*aw→fe|n*	*softe|ne|d*		
17.5	*oi*	*o0,i2,y1*	*boil* *point*	*chamois* *choir* *heroin* *jingoiSm* *memoir* *porpoiSe|*	
17.5.1	*go|ing*	*g,o1,w,i1,ng,g*	*fOre|go|ing*		
17.6.1	*ol[dt]*	*→Ol[dt]*	*fold* *soldie|r* *bolt*	*harold* *solde|r*	
17.6.2	*olk*	*o0,k*	*folklOre|* *polka*		
17.6.3	*rol*	*→rOl*	*patrol* *prolif|ic* *roll*	*frol|ic* *petrol* *rollick* *trolle|y*	*carol* 2.10.1
17.6.3.1	*rolog*	*r→olog*	*astrology*		
17.7.1	*o[mn]#*	*u1,[mn]*	*from* *kingdom* *nAtion* *washington*	*tom* *don*	
17.7.2	*o[mn]s#*	*u1→[mn]s#*			
17.8	*ong*	*aw→ng*	*long* *strong|ly*	*congressional* *sponge|* *tongue|*	
17.9	*on't*	*→Ont*	*don't* *won't*		
17.10	*oo*	*iu,iu*	*brood* *coope|r* *pool*	*brooch* *cooperAte|* *soot* *wool*	

	Rule		Examples	Exceptions	Compare
17.10.1	*ook*	*oo,k*	*took* *book*	*spook*	
17.10.2	*o|or*	*→oor*	*po|or*		
17.10.2.1	*do|or*	*→dor*	*outdo|ors*		*doorway*
17.10.2.2	*flo|or*	*→flor*	*subflo|or*		*floorboard*
17.10.3	*coord*	*→cOord*	*coordinAte|*		
17.10.4	*foot*	*f,oo,t*	*afoot* *footage|*		
17.10.5	*[ghw]ood*	*[ghw],ou,eu→d*	*good|ness* *neighborhood* *woode|d*	*hoodoo*	
17.10.6	*lood*	*→lud*	*blood* *flood*		
17.10.7	*stood*	*s,t,ou,eu,d*	*understood*		
17.11	*or*	*o1,r*	*for* *morbid*	*AviAt|or†*	
17.11.1	*orough*	*→urrow*	*borough* *thorough|ly*		
17.11.2	*wor*	*w→er*	*word* *worry*	*sword*	
17.12	*oss#*	*aw→ss*	*cross*	*gross*	*cossack* *mossy*
17.13	*other*	*u2,u2,dh,u3,er*	*otherwISe|*		*OTHe|r* 16.1
17.14	*ou*	*→ow*	*doubt* *rout* *our*	*colour* *couSin* *mourn* *ouzel* *poultice|* *rouge|* *southern* *tour*	*routIne|* 9.12
17.14.1	*ou#*	*→oo*	*caribou*		
17.14.2	*oubl*	*u1,b,l*	*doublet* *trouble|*		
17.14.3	*ough*	*u1→f*	*enough* *tough*	*bough* *cough*	*thorough* 17.11.1
17.14.3.1	*ought*	*aw,t*	*bought* *thought*		
17.14.3.2	*though#*	*→THo#*	*although*		*thought*
17.14.3.3	*through*	*th,r,iu,iu*	*throughout*		
17.14.4	*ould*	*ou,eu,d*	*could* *wouldn't*		
17.14.4.1	*oulde|*	*→Olde|*	*boulde|r* *shoulde|r*		
17.14.5	*oup*	*iu,iu,p*	*group* *troupe|*	*cantaloupe|*	
17.14.5.1	*coupl*	*k,u1,p,u2,l*	*coupl|ing*		*coupe|*
17.14.6.1	*ous#*	*u2→s#*	*vArious*		*roustabout*
17.14.6.2	*ous|ly*	*u2→sly*	*marvelous|ly*		*hous|ing*
17.14.6.3	*ous|ness*	*u2→sness*	*nervous|ness*		

	Rule		Examples	Exceptions	Compare
17.14.7	*countr*	*→kuntr*	*country*		*counte|r*
17.14.8.1	*ource*	*→orce*	*reSource|*		
17.14.8.2	*cour*	*→kor*	*course|* *court*	*courage|* *scour*	
17.14.8.3	*#four*	*f,o1,w,er*	*fourth*		
17.14.8.4	*journ*	*→jern*	*adjourn* *journe|y*		
17.14.9	*touch*	*→tuch*	*untouch|able*	*cartouche|* *touché*	
17.14.10	*you*	*y2,iu,iu*	*your* *youth*		
17.14.10.1	*young*	*y2→ung*	*youngste|r*		
17.15	*#over*	*→Over*	*overthrow*		
17.16	*ow*	*a0,w*	*brow* *howl* *towe|r*	*mow* *stow*	
17.16.1	*#ow*	*o2,o2,w*	*owe|* *owne|d*	*owl*	
17.16.2	*know*	*n,o1,o1,w*	*unknown*		
17.16.2.1	*knowledg*	*→nowledg*	*knowledge|able*		
17.16.3	*low*	*l,o0,w*	*blowe|r* *flow* *shallow*	*glowe|r* *plow*	
17.16.3.1	*#allow*	*u2,l→ow*	*allowance|*		
17.16.3.2	*flowe|r*	*f,l→ower*	*flowe|ry*		
17.16.4	*nowledg*	*n,a0,l,e2→j*	*acknowledg|ment* 4.3		
17.16.5.1	*#row*	*r,o0,w*	*row* *rowe|d*	*rowdy*	*brow* *crow* *trowel*
17.16.5.2.1	*arrow*	*ae,r,o0,w*	*harrow|ing* *narrow*		
17.16.5.2.2	*orrow*	*a1,r,o0,w*	*borrow*		
17.16.5.2.3	*urrow*	*u1,r,o0,w*	*furrow* *thorough* 17.11.1		
17.16.5.3	*grow*	*g,r,o0,w*	*ingrown*	*growl*	
17.16.5.4	*throw*	*th,r,o0,w*	*overthrown*		
17.16.6	*show*	*sh,o0,w*	*show|ing*	*showe|r*	
17.16.7	*snow*	*s,n,o0,w*	*snowe|d*		
17.16.8	*toward*	*→tord*	*untoward*		
17.17	*oy*	*→oi*	*boy* *oyste|r*	*coyOte|*	
17.18	*logy*	*l,u2→gy*	*bIology* *trilogy*		*lOgy*
17.1x	*fOre|most* *most|ly*				
17.20	*phot*	*→fOt*	*photocopy* *photon*		
17.21.1	*post#*	*→pOst*	*impost*		*expostUlAte|* *imposte|r*

	Rule		Examples	Exceptions	Compare
17.21.2	*#post*	*→pOst*	*postage|* *poste|d*	*postUlAte|*	
17.22	*togETHe|r*	*t,iu,iu,g,e1,dh,er*	*altogETHe|r*		
18	*p*	*p*	*pept|ic* *stop*	*corps* *coup* *cupboard* *raspberry†* *receipt*	
18.1	*pp*	*→p*	*grapple|* *sapphIre|*		
18.2	*ph*	*f*	*photograph*	*haphazard* *phthalein* *shepherd*	
18.3	*#p[nst]*	*→[nst]*	*pneumat|ic* *pseudonym* *ptarmigan*		
19	*q*	*k*	*iraq*		
19.1	*qu*	*q→w*	*quarrel* 2.22 *requIre|* *squeeze|*	*coquette|* *liqu|or* *mosquito*	
19.1.1	*que|*	*→ke|*	*cheque|* *torque|*		*oblique|* 10.6 *risqué* 7.3
19.1.1.1	*que|st*	*→quest*	*reque|st*		
20	*r*	*r*	*rAre|ly*		
20.1	*rr*	*→r*	*dIarrhea* *horr|or*		
20.2	*rh*	*→r*	*rhythm*		*neighborhood* 17.11 *perhaps*
21	*S*	*→z*	*nOSe|* *pleaSe|* *viSible|*	*ceaSe|* *dOSe|*	
21.1	*Si[aOo]*	*zh,ih,u2*	*ASia* *enviSiOne|d* *confUSion*		*russia* 22.8
21.2	*Sm*	*z,u3,m*	*priSm*	*priSmat|ic†* *taliSman†*	
21.3	*Sua*	*zh→Ua*	*caSualty*		
22	*s*	*s,s*	*cost* *star*	*apropos* *debris* *demesne|* *Isle* *viscount*	
22.1	*ss*	*→s*	*passion* *rest|less|ly*	*sciss|ors†*	
22.2	*s#*	*z,s*	*his* *bans*	*ant|ics†* *puffs†* *ticks†* *pops†* *cats†* *plus* *marquis*	
22.3	*ss#*	*s,s*	*less*		

	Rule		Examples	Exceptions	Compare
22.4	*sc[Eei]*	*→c[Eei]*	*scEne|ry* *coalesce|* *discern* *conscience|* *scimitar*		*scIon*
22.4.1	*#scien*	*s,a0,iy,e2,n*	*scientif|ic*		*conscience|* *antiscience|*
22.5	*sch*	*sh*	*fische|r*	*unschoole|d*	
22.5.1	*#sch*	*s,k*	*schedUle|* *scholar*	*schiSm* *schist*	
22.6	*scle|*	*→sle|*	*corpuscle|* *muscle|*		*sclerot|ic*
22.7	*sh*	*sh*	*shush*	*dishearte|n* *mishap* *sheepshead*	
22.8	*si[aOo]*	*sh,ih,u2*	*russian* *passion* *pensiOne|r*		*viSion* 22.8
22.9	*sy#*	*z,z,y1*	*pansy*		
22.9.1	*ssy#*	*s,s,y1*	*grassy*		
23	*t*	*t*	*toaste|r*	*mortgage|*	*ofte|n* 17.4.1
23.1	*t#*	*t,h*	*bat*	*depot*	
23.2	*te|#*	*t,h*	*abAte|*		
23.3	*tt*	*→t*	*butte|r*	*cattail†*	
23.3.1	*ttl*	*t,u3,l*	*little|*		
23.4	*TH*	*dh*	*bATHe|* *clOTH|ing*	*auTH|or†*	
23.5	*th*	*th*	*bath* *theory*	*asthma* *chatham* *hOthead* *northern†* *posthUmous* *thomas* *worthy†*	
23.5.1	*#there*	*dh,ai,y1,r*	*therewith*	*theremin*	
23.5.2	*thm*	*dh,u3,m*	*logarithm* *rhythm*	*arithmet|ic*	
23.5.3	*with*	*w,ih,dh*	*forthwith* *withheld*	*beckwith†*	
23.6.1	*tiA*	*sh,ih→A*	*initiAte|*		
23.6.2	*tia*	*sh,ih,u2*	*hermitian* *partial*		
23.6.2.1	*sti[ao]*	*s,t,ch,ih,u2*	*christian* *combustion*		
23.6.3	*tien*	*sh,ih,e2,n*	*pAtient*		
23.6.4	*tio*	*sh,ih→o*	*contentious* *rAtio* *stAtion*	*cAtion* *conventional†* *equAtion†*	
23.6.4.1	*tiOn*	*→tion*	*mentiOne|d*		
23.7	*ts* (for hardware bug)	*t,t→s*	*hits*		

	Rule		Examples	Exceptions	Compare
23.8	*tu[aeo]*	*t,ch→u[aeo]*	*eventual* *constituent* *virtuous* *virtue|*		
23.9	*ste|n*	*→se|n*	*faste|n* *liste|n*		
23.10	*stl*	*→sl*	*castle|* *whistl|ing*		*last|ly*
24	*U*	*y2,iu*	*distribUte|* *execUtIve|*	*lettUce|*	
24.1	*#Un*	*→un*	*Un|able* *UnEve|n*	*Unanimous*	
24.1.1	*#Un[Ii]*	*→Un[Ii]*	*Union* *Unit* *UnIte|d*	*Uninforme|d*	
24.2	*bUry*	*→berry*	*canterbUry*		
24.3	*#bUSi*	*b,i1,z,z,i2*	*bUSi|ly* *bUSi|ness*		
24.4	*[dlrst]U*	*[dlrst],iu,iu*	*dUbious* *salUte|* *rUle|* *sUSan* *stUdent*	*edUcAtion†* *modUle|†* *lUrid†*	*dew* 7.15.1
24.4.1	*dUr*	*d→Ur*	*verdUre|*	*dUr|able†*	
24.4.2	*lUre*	*l→Ure*	*failUre|*	*lUre|†*	*silUrian*
24.4.3	*sUr*	*sh,sh,iu,iu→r*	*sUre|* *insUrance|*		*cloSUre|* 31.3 *pleaSUre|* 7.4.2
24.4.4	*tUl*	*t,ch,y2,yu.l*	*petUlant*	*tUlarEmia*	
24.4.5	*tUr*	*t,ch,iu,r*	*natUral* *overtUre|*		
24.5	*jU*	*d,j→U*	*jUdicious*	*jUry†*	*jewel* 17.15.1
24.6.1	*stUdie|*	*→studie|*	*stUdie|d*		*stUdious*
24.6.2	*stUdy*	*→study*	*understUdy*		
25	*u*	*u0*	*hubbub* *puddle|*	*pudd|ing*	
25.1	*u#*	*iu,iu*	*malibu* *tUtu*	*menu*	*mInus* *tUtus*
25.2	*u|*	*→U|*	*continu|ing* *valu|able*	*intrigu|ing*	
25.3	*u[Aao]*	*→U[Aao]*	*graduAte|* *manual* *continuous*		
25.3.1	*gua*	*→gwa*	*guano*		*language|* 15.3.3
25.3.1.1	*guar*	*→gar*	*guarantee|* *guard*		
25.3.2	*suA*	*s,w→A*	*persuAde|*	*suAve|*	
25.3.3	*Sua*	*zh→Ua*	*caSual* *viSual*		
25.4	*ue*	*iu,iu→e*	*duet*	*bluebird* *guerdon*	*suEde|*

	Rule		Examples	Exceptions	Compare
25.4.1	*ue|*	*→Ue|*	*continue|d* *venue|*	*pursue|d†*	*blUe|*
25.4.1.1	*gue|*	*g→e|*	*catalogue|* *fUgue|* *league|*		*intrigue|* 10.6
25.4.1.1.1	*#argue*	*→argUe*	*argue|d*		*deSargue|s*
25.4.2	*guess*	*g→ess*	*guesse|s*		
25.5	*ul*	*iu,iu*	*brulSe|* *recrult|ing* *sult|able*		*requlte|* 19.1
25.5.1	*uln*	*→Uin*	*genulne|* *rulne|d*		
25.5.2	*gul*	*g→l*	*gulde|* *gulle|* *disgulSe|*	*begulne|*	
25.6	*ui*	*iu,iu,i1*	*continuity* *ruin*	*guinea*	*continu|ing* 25.2
25.6.1	*uil*	*i1,l*	*build* *guilty*		
25.6.2	*cuit*	*k,i1,t*	*circuit*	*circuitous*	
25.6.3	*guis*	*g,i1→s*	*roguish*		*anguish* 15.3.3
25.6.4	*[rs]uit*	*→[rs]Ut*	*fruit* *pursuit*	*fruition*	*jeSuit*
25.7	*uo*	*→Uo*	*contiguous* *virtuous*		*vacuOle|*
25.8	*ur*	*u2,er*	*purple|* *urgent*		
25.9	*[bp]ush*	*[bp],ou,eu,sh*	*ambush* *pushcart*		*cushion* *plush*
25.10	*uu*	*→Uu*	*vacuum*		
25.11	*uy*	*→I*	*buye|r*		
25.12	*ful*	*f,ou→l*	*help|ful* *ful|ly*		
25.13	*#pull*	*p,ou→l*	*pulle|y*		*pulsAte|*
25.14.1	*put#*	*p,oo,t*	*output*		*putty*
25.14.2	*puts#*	*p,oo→ts#*			
25.15	*truth*	*t,r,iu,iu,th*	*truth|ful*		*ruth|less* *rutherford*
26	*v*	*v*	*valve|*	*molotov†*	
27	*w*	*w*	*boxwood* *wet*	*sword*	
27.1	*wh*	*h,w*	*what* *when*	*whoop* *whOre|*	*blowhard* 17.16.3
27.1.1	*wh[Oo]l*	*→hOl*	*whOle|hearte|d|ly* *whol|ly*		
27.2	*wr*	*→r*	*wrIte|*		
27.3	*answe*	*→anse*	*answe|r*		*unswerv|ing*
28	*x*	*→ks*	*box* *exit*	*exist†* *exUde|†*	
28.1	*#x*	*→z*	*xenon*		