Statistics of Operationally Defined Homonyms of Elementary Words*

by L. L. Earl, B. V. Bhimani, and R. P. Mitchell
Lockheed Palo Alto Research Laboratory, Palo Alto, California

This computerized study of the homonyms of elementary words (roughly equivalent to monosyllabic words) has allowed the compilation of exhaustive lists of homonym sets, using phonetic transcriptions from five different dictionaries. Of the 5,757 elementary words, 2,966 were involved in at least one homonym set, indicating that homonyms will present a significant problem in mechanized word recognition. The effects on the homonym sets of changing from the phonetic transcription of one dictionary to another were tabulated, as were the effects of removing dialectal pronunciations. Since the effects of dialectal variations turned out to be relatively small, it was possible to categorize and list for study the actual words whose dialectal pronunciations caused homonym-type confusion with other words.

Introduction

In 1919 Robert Bridges published an essay on homonyms as Tract II of the Society for Pure English in which he compiled lists of words that are pronounced alike but have "different origin and signification." His lists, drawn from the entire language, contained 835 entries comprising 1,775 words, which led him to the propositions that homonyms are a nuisance and that English is exceptionally burdened with them. He proposed also that homonyms are self-destructive and tend to become obsolete, a proposition which may be questioned in the light of the number of homonyms discovered in our investigations.

Words that are pronounced the same but have different spellings and meanings, variously called either "homonyms" or "homophones," are of even more practical interest today than in 1919, because automatic handling of spoken languages will require distinguishing among them. Our results indicate that over half the one-syllable words in English are homonyms according to at least one dictionary, showing certainly that homonyms are a significant class of words. Because we have been able to use automatic processing in working with more than one dictionary, we believe our studies are also helpful in providing insight into phonetic transcription systems.

Method of Compilation

We have undertaken an exhaustive compilation of homonym sets among elementary words from five dictionaries which give phonetic transcriptions. A homonym set is defined here as a set of different orthographic forms having an identical phonetic transcription in a specified dictionary. We did not investigate

either meaning or origin. Any member of a homonym set is called a "homonym." Elementary words, defined by J. L. Dolby and H. L. Resnikoff,[1] are roughly equivalent to one-syllable words, differing only because of simplifications made in the recognition of one-syllable words from the orthographic form. (For example, a final *e* was not regarded as a syllabic vowel except under special circumstances, and as a consequence, a small set of words like *he, be, we,* etc., are not included in elementary words although they are one-syllable words.) The elementary words provide a set of words sufficiently small so that it is practical to undertake an exhaustive automatic compilation, yet they are a particularly significant set for two reasons:

(1) the frequency of occurrence of homonyms is much greater in elementary than in multisyllable words; and

(2) most of the occurring variations in syllabic spelling show up in elementary words.

The five dictionaries[2-6] used in this study will be referred to by the following abbreviations.

MW3—*Webster's Third New International Dictionary of the English Language;*

KK— A *Pronouncing Dictionary of American English,* by Kenyon and Knott;

ACD— *The American College Dictionary;*

JON— *Everyman's English Pronouncing Dictionary,* by Daniel Jones;

SOX— *The Shorter Oxford Dictionary on Historical Principles.*

SOX and JON represent speech patterns in Great Britain; sometimes variant British pronunciations are given in JON. The other three dictionaries represent speech patterns in the United States: ACD represents the midwestern speech pattern, with occasional variant pronunciations given; KK presents separately the pronunciation of words in eastern, southern, and midwestern "dialects"; and MW3 presents speech in re-

gions considered by KK and also in regions of New York City (e.g., Brooklyn and the Bronx) and in regions of the south where the "el" sound is dropped.

The homonyms were derived separately for each dictionary, so that differences in the phonetic symbology of the dictionaries did not cause any problems. For each compilation, all 5,757 elementary words were considered, even though each word did not appear in all five dictionaries. (For missing words, probable pronunciations were used, suitably marked, as will be explained.) The homonym sets were derived automatically from the dictionaries on magnetic tape. In these tape dictionaries each word appeared in its graphic form, split into consonant and vowel strings, with its phonetic transcription in code. A word with more than one pronunciation occurred more than once. Each occurrence of the word was identified by dictionary source and by class of dialect when applicable. Thus for ACD, ACD1 indicated the standard midwestern pronunciation, and ACD2 a variant. Table 1 gives the meanings of all the codes used. Markers were added to these codes to identify special cases of phonetic transcriptions, which arose as follows.

TABLE 1
PHONETIC REPRESENTATION CODES

Code	Interpretation	Dictionary
JON 1 ...	First pronunciation	JON
JON 2 ...	Second pronunciation	JON
ACD 1 ...	First pronunciation	ACD
ACD 2 ...	Second pronunciation	ACD
101SK ...	Midwestern pronunciation	KK
102SK ...	First variant pronunciation	KK
103SK ...	East and South pronunciation	KK
104SK ...	East pronunciation	KK
105SK ...	Second variant pronunciation	KK
106SK ...	Third variant pronunciation	KK
107SK ...	Fourth variant pronunciation	KK
101SW ...	Midwestern pronunciation	MW3
102SW ...	First variant pronunciation	MW3
103SW ...	Boston R-dropper pronunciation	MW3
104SW ...	Brooklyn R-dropper pronunciation	MW3
105SW ...	L-dropper pronunciation	MW3
106SW ...	Second variant pronunciation	MW3
107SW ...	Third variant pronunciation	MW3
108SW ...	Fourth variant pronunciation	MW3
109SW ...	Fifth variant pronunciation	MW3
20XSW ...	Consonant variant pronunciation on the 10X pronunciation of	MW3
20XKK ...	Consonant variant pronunciation on the 10X pronunciation of	KK

Instead of transcribing phonetics from the dictionaries, an algorithm (about 93 per cent accurate) was used which automatically generated the phonetic form or forms for each dictionary from the graphic form. The generated forms were manually checked three times against the dictionaries, and errors were corrected. Corrected words were marked with a D indi-

cator, for example, the code 101DK is equivalent to 101SK, except that this pronunciation was not derived algorithmically. The phonetic representations of words missing from a given dictionary could not be directly checked, however, and were marked with an N indicator if the algorithm had functioned correctly in deriving the SOX phonetics of that word, or an M indicator if the algorithm had given incorrect results on the SOX dictionary, in which case the probable error had been corrected. Thus, the M indicator is almost equivalent to an N + D marker. The algorithms for generating phonetic transcriptions and the correction procedures are completely described in an unpublished manuscript by Bhimani and Mitchell.[7]

Phonetic transcriptions were generated by algorithm because the homonym study grew out of the more general study described,[7] and was designed to meet its requirements. To make a meaningful study of the relationship between orthographic and phonetic forms, it seemed desirable to work with the entire set of data available in the dictionaries chosen. Since there is quite a discrepancy among the dictionaries in the words listed, and in the dialect pronunciations given for words, the algorithmic method of deriving the phonetic codes is the only one in which all the words can be utilized. (If only words common to all dictionaries are used, the data set is cut roughly in half.) Also, the algorithmic method is easier in that it is difficult for keypunchers to interpret the phonetic markings of a dictionary. Thus, keypunching would be expensive, and many more corrections would be necessary. Since the generated forms were carefully checked, no bias will have been introduced by using the algorithm for phonetic forms which are spelled out by the dictionaries. Also, since the algorithm shows a 93 per cent accuracy in assigning phonetic codes which can be checked with the dictionary, it is reasonable to expect that the use of phonetic codes which cannot be checked will not introduce more than about a 7 per cent error. (Actually, the error can be expected to be less than 7 per cent in view of the elaborate checking and comparing programs which were used.[7]

Once the words with their phonetic transcriptions and dictionary codes were on tape in the format just described, homonym compilation was merely a matter of sorting or grouping words with the same phonetic transcriptions. Figure 1 shows part of a page from one of the homonym printouts. The first three columns give the graphic form split into consonant and vowel strings; the next three columns give the code for the phonetic representation; and in the final column, the numbers indicate the dialect represented, and the letters indicate the dictionary source (in this figure, Kenyon and Knott[3]) and the algorithmic derivation of the phonetic representations. A blank line separates the homonym sets.

Fig. 1.—Sample entries of homonym printout

Discussion of Results

The number of sets and number of total words involved in homonym sets differ considerably from dictionary to dictionary, and a word may be in a homonym set according to one dictionary's phonetic representation but not according to another. The statistics of the homonym sets in each of the five dictionaries are given in Table 2 and Figure 2. (Note the 10 to 1

TABLE 2
NUMBER OF HOMONYM SETS IN FIVE DICTIONARIES

NUMBER OF WORDS IN A SET	TOTAL NUMBER OF SETS				
	MW3	KK	ACD	JON	SOX
2	1,889	1,402	717	727	661
3	380	268	133	142	117
4	99	55	33	31	27
5	18	11	4	8	3
6	9	5	2	0	0
7	1	1	0	0	0
8	1	0	1	1	0
9	0	1	0	0	0
10	1	0	0	0	0

change in scale in Fig. 2 between sets of three and sets of four.)

When the discrepancies among dictionaries turned up, a program was written to show for each word which phonetic transcriptions gave rise to homonym sets. Figure 3 is a sample page of the output (hereafter called the "homonym comparison tables") from this program. It indicates that the word *fon* is involved in a homonym set only according to the standard MW3 pronunciation, yet the word *forte is* involved in six MW3 homonym sets, four KK sets, one JON set,

one ACD set, and no SOX set. In general, SOX has the fewest homonyms, indicating perhaps that the SOX phonetic transcription is finer. Of course SOX gives only one pronunciation while the others give variants, which will reduce the number of homonyms for SOX. Still, there appear to be quite a few words for which the JON1, ACD, 101SK, and 101SW pronunciations all give rise to homonyms while the SOX pronunciation does not. The total number of words in the homonym comparison table is 2,966, showing that 2,966 of the 5,757 elementary words are in a homonym set according to at least one dictionary. Thus, the homonym comparison table shows that over 50 per cent of the elementary words can be considered ambiguous in their spoken form. For about 50 per cent of these words, there is disparity among the dictionaries in homonym membership.

Before exploring the possible reasons for the disparity in homonym sets, some possibilities can be eliminated. Since these dictionaries were published at approximately the same time, and since it is generally recognized that their contents are periodically updated, historic vowel changes are not expected to cause discrepancies. Also, vowels which are consistently pronounced one way according to one dictionary, and another way (but always the same other way) according to a second dictionary, will affect the homonym compilation very little. For example, *break* and *brake* are homonyms whether the vowel is given a British pronunciation as indicated by "b r e i k" in JON or an American pronunciation as indicated by "b r e k" in KK. The following list gives the phonetic symbols for this sound from each of the five dictionaries and the corresponding code used for machine purposes. (JON and KK use the International Phonetic Alphabet.)

SOX	bre'k	BRE1419K
JON	breik	BREIK
ACD	brāk	BRA4K
KK	brek	BREK
MW3	brāk	BRA4K

Thus, consistent changes from dialect to dialect will not cause significant discrepancies in homonyms.

Variant spellings given in some dictionaries will result in "extra" homonyms from a semantic point of view. Such "extra" homonyms do not, however, account for discrepancies among dictionaries because all of the words were used in the study of each dictionary, and the same extra homonyms would be expected in each compilation. Moreover, variant spellings were noticed during the three manual checks of the dictionaries, but their number seemed so small that it was not considered serious enough to warrant isolation.

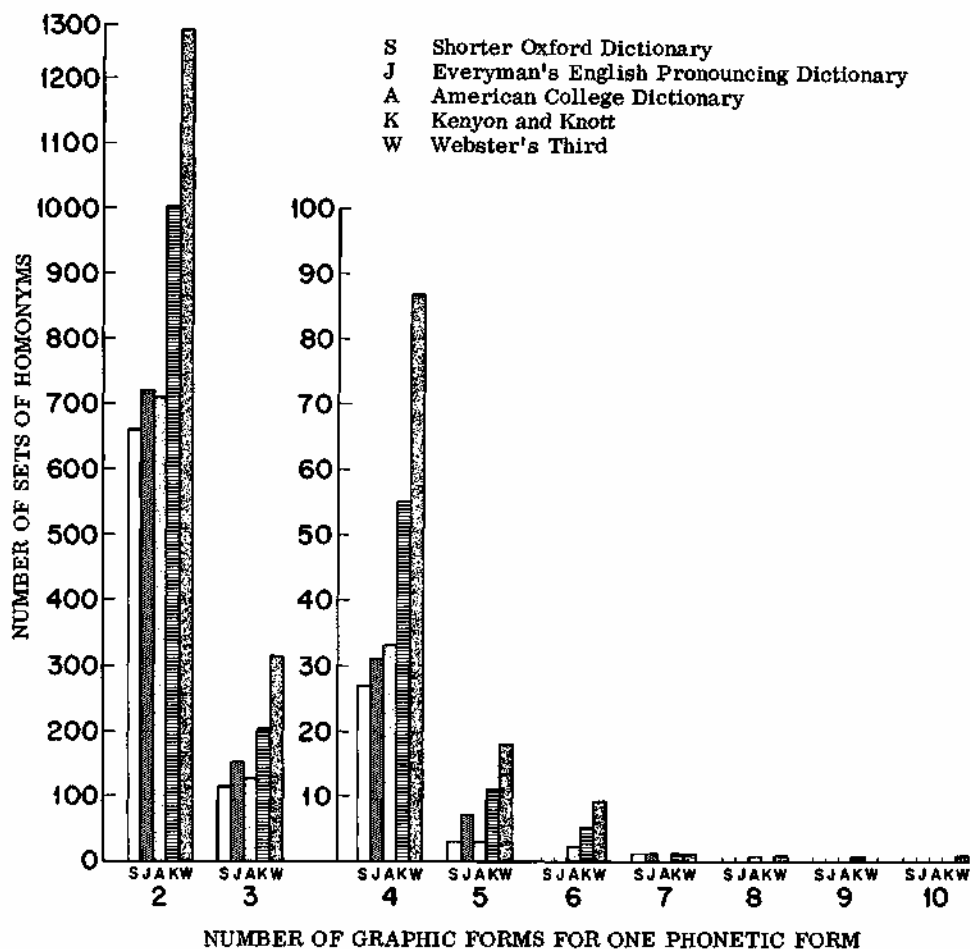What then will cause discrepancies from dictionary

Figure 2

1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0

NUMBER OF SETS OF HOMONYMS

S   Shorter Oxford Dictionary
J   Everyman's English Pronouncing Dictionary
A   American College Dictionary
K   Kenyon and Knott
W   Webster's Third

100
90
80
70
60
50
40
30
20
10
0

S J A K W   S J A K W   S J A K W   S J A K W   S J A K W   S J A K W   S J A K W   S J A K W   S J A K W
2   3   4   5   6   7   8   9   10

NUMBER OF GRAPHIC FORMS FOR ONE PHONETIC FORM

FIG. 2.—Graphic presentation of number of homonym sets in five dictionaries

FIG. 3.—Entries from the homonym comparison table

to dictionary? When several dialects are considered together in the compilation of homonyms, as in KK and MW3, extra homonym sets or larger sets can be produced across the dialects. For instance, two words which are not homonyms within either dialect A or dialect B may become homonyms when the dialect A pronunciation of one is compared with the dialect B pronunciation of the other. Thus *rear* and *rare* have different pronunciations if only the midwestern and first variant pronunciations are compared, but the second variant pronunciation of *rear* is identical to the eastern pronunciation of *rare*. By removing the dialect pronunciations from the homonym sets, two objectives are met: (1) the ambiguity producing effects of dialects are shown, and (2) homonym disparities between ACD and KK or MW3 which result from the inclusion of dialects are removed.

In removing dialects, some difficulty is encountered in identifying true dialectal pronunciations. The 103SK, 104SK, 20XSK (where X is any number), 103SW, 104SW, 105SW, 30XSW, and 20XSW pronunciations (Table 2) were considered to be true dialects by the dictionaries in which presented and were, therefore, removed by computer program from the homonym sets. The homonym comparison program was run again on the homonyms after the removal of the dialectal pronunciations to produce another com-

parison table of the same form as shown in Figure 3. The results show the expected reduction in the number of sets containing a given word and in the number of words that appear in homonym sets, but these reductions are not so large as was expected.

TABLE 3

STATISTICAL SUMMARY OF WORDS INVOLVED IN HOMONYM SETS, SHOWING EFFECT OF DIALECT REMOVAL

SET DESCRIPTION (TOTAL SET)	NUMBER OF WORDS IN SET	
	With Dialects	Without Dialects
Words forming a homonym in at least one dictionary	2,966	2,714
Words forming a homonym in one dictionary	746	535
Words forming a homonym in two dictionaries	236	214
Words forming a homonym in three dictionaries	189	184
Words forming a homonym in four dictionaries	290	297
Words forming a homonym in all dictionaries	1,505	1,484
Words forming a homonym in SOX	1,754	1,743
Words forming a homonym in ACD	1,937	1,937
Words forming a homonym in JON	2,039	2,039
Words forming a homonym in MW3	2,600	2,297
Words forming a homonym in KK	2,140	2,096

The homonym comparison tables were used to compile some statistics of homonym membership, to show the relationships among the dictionaries. These statistics, compiled both before and after the removal of dialects, are shown in Table 3. Note that with the dialects removed, the number of elementary words which are in homonym sets is reduced only about 5 per cent, from 52 to about 47 per cent. Note also that the relationships among the various sets named in Table 3 do not change significantly. In particular, the ratio between the words forming a homonym in all dictionaries and the words forming a homonym in any dictionary changes only from 0.5074 to 0.5467 when dialects are removed. Thus, the dialects are not the main reason for the large number of homonyms, nor are they the major cause of discrepancies among the dictionaries.

It is also revealing to consider the actual occurrence of ambiguity introduced by the dialects, and because they are not numerous we have prepared tables which give them all. In Table 4, Part A shows all new sets introduced by the dialect pronunciations of KK; Part B shows all words or sets added to nondialectal homonym sets by a dialect pronunciation of KK. The starred items were not removed by the program but seemed to the authors to be dialect forms and were removed later.

TABLE 4

WORDS INVOLVED IN HOMONYM SETS IN KK BECAUSE OF DIALECTAL PRONUNCIATIONS

PART A

Graphic	Phonetic	Dictionary Code
MUZZ	MA6Z	101NK
MUS		201NK
DAZE	DEZ	101SK
DASE		201NK
GREITH	GREPI	201NK
GRAETH		101DK
NAIS	NEZ	201NK
NAZE		101NK
CLEAR	KLE1E2(R	105SK
CLARE		104SK
REAR	RE1E2(R	105SK
RARE		104SK
MY	ME2	103 or 104SK
	or	
MAC	MI	105 or 106SK
BROOSE	IUIZ	202NK
BRUISE		102DK
CHESE	T$I1Z	201NK
CHEESE		101NK
CROZE	KROZ	101NK
CROSE		201NK
SHORE	$0E2(R	104 or 103SK
SURE		105 or 102SK
HAUSE	HOIZ	201NK
HAWSE		101NK
BROOSE	BRUIZ	201NK
BRUISE		101DK
COUTH	KUIPI	101SK
COOTH		201MK
JEER	DZ1E1E2(R	105SK
*GEER	DZ1E1E2(R	105SK
JEER	DZ1E1E2(R	105SK
FEAR	FE1E2(R	105SK
*FEER	FE1E2(R	105SK
FLEAR	FLE1E2(R	105SK
*FLEER	FLE1E2(R	105SK
HEAR	HE1E2(R	107SK
*HEER	HE1E2(R	107SK
HERE	HE1E2(R	108SK
LEAR	LE1E2(R	105SK
*LEER	LE1E2(R	105SK
TEAR	TE1E2(R	106SK
*TEER	TE1E2(R	105SK
TIER	TE1E2(R	105SK
WEIR	WE1E2(R	105SK
*WERE	WE1E2(R	105SK
TROTH	TRA3PI	105SK
*TROUGH		106DK
BUM	BA6M	101SK
*BOMB		102SK

EARL, BHIMANI, AND MITCHELL

Part B		
NEEZE	NIIZ	101NK
°WERE	WA1E2(R	107SK
°OUR	A2U(R	105SK
°EAR	E1E2(R	105SK
BIER	BE1E2(R	105SK
°BEER	BE1E2(R	105SK
°BLEAR	BLE1E2(R	105SK
DEER	DE1E2(R	105SK
°DEAR	DE1E2(R	105SK
°KIER	KE1E2(R	105SK
°MEER	ME1E2(R	105SK
°PEER	PE1E2(R	105SK
SPEAR	SPE1E2(R	105SK
°SPEERE	SPE1E2(R	105SK
°CHEER	T$E1E2(R	105SK
°AND	E2N	106SK
°WEAR	WIE2(R	106SK
°POOR	POE2(R	105SK
°PRYSE	PRAIZ	201SK
°BLOUSE	BLAUZ	201SK
°CLOUGH	KLA2F	103DK
°DON	DA3N	103SK
°WOT	WA3T	103SK
°SHARE	$E1E2(R	103SK
°CERE	SE1E2(R	104SK
°ERR	E3(R	103SK
°YAIR	JE1E2(R	104NK

Table 5 (pages 24 and 25) shows all the dialectal pronunciations removed from MW3, but here we have divided them into nine significant categories as follows:

Set A.—New homonym sets in which a pronunciation of type 20X (where again X is any number) is involved. These reflect confusion between T and D or S and Z sounds, which may not be strictly a dialectal phenomenon.

Set B.—New homonym sets in which a pronunciation of the type 20X is not involved.

Set C.—Words in which a pronunciation of the type 20X adds one to the number of homonyms in a non-dialectal homonym set.

Set D.—Same as C, except a non-20X dialectal pronunciation is responsible for an extra member of a homonym set. (Starred items were added by hand, as in Table 6-4.)

Set E.—New homonym sets caused by a pronunciation of the type 20X, where each of these sets has the same pronunciation as a non-dialectal homonym set.

Thus, these words add more than one member to a non-dialectal set.

Set F.—Same as E, except a non-20X dialectal pronunciation is responsible for the extra members to homonym sets.

Set G.—Words in which a dialectal pronunciation causes confusion with words already in sets B or D. Thus, a dialectal pronunciation of *chert* causes the homonym set *chert, chat*. A dialectal pronunciation of *chad* adds to the set, making it *chert, chat, chad.*

Set H.—New homonym sets in which two dialectal variations combine to form a homonym group.

Set I. —New homonym sets in which two dialectal variations combine to form a homonym group, where each of these groups has the same pronunciation as a non-dialectal homonym set.

Summary and Conclusions

To summarize our results, an exhaustive compilation of the homonyms of elementary words shows that a surprisingly high percentage of these words (30 per cent at the best, more than 50 per cent at the worst) are homonyms. Furthermore, considerable discrepancy in the homonym data among the five dictionaries used has been made apparent. Neither of these results changed significantly with the removal of the dictionary-defined dialectal vowel variations. The latest tests show that limiting the words considered in compiling homonyms to those with standard meanings in both SOX and MW3 does help somewhat to even out the discrepancies, at least among the three dictionaries KK, JON, and ACD. Statistical results of homonyms among double standard words are given in Table 6.

TABLE 6

NUMBER OF HOMONYM SETS AMONG
DOUBLE STANDARD WORDS

NUMBER OF WORDS IN A SET	TOTAL NUMBER OF SETS				
	MW3	KK	ACD	JON	SOX
2	709	591	578	590	311
3	102	87	66	86	31
4	21	12	13	9	6
5	1	1	0	1	0
6	2	0	0	0	0
7 or more	0	1	1	1	0

Obviously we have not yet really accounted for the discrepancies. Also, though reducing the size of the data set inevitably reduces the number of homonyms, even in this data set of non-specialized, non-foreign, and non-archaic words, the homonyms make up a significant percentage of the words, and there is a large number of phonetic ambiguities with which mechanized word recognition must deal.

TABLE 5

Words Involved in Homonym Sets in MW3 Because of Dialectal Pronunciations

Set A

PUT	DROWTE	SATE	SNOT	WET	CLEAT	LIT	QUOTE
PUD	DRAD	SADE	SNOD	WED	CLEAD	LID	QUOD
NEWT	CLOUT	SLATE	TROT	CHUT	LEASE	MITT	TOTE
NUDE	CLOUD	SLADE	TROD	CHAD	LEESE	MOD	TOAD
FAT	CROUT	TRAIT	BET	GLUT	PLEAT	WRIT	BROUGHT
FAD	CROWD	TRADE	BED	BLOOD	PLEAD	ROD	BROAD
GAT	LOUT	BRAT	BRUTTE	HUT	SPETE	SKIT	BRAUGHT
GAD	LOUD	BROD	BUD	HUD	SPEED	SKID	BRAUD
HAT	BLATE	DOT	FET	CRUT	TWEET	FRIGHT	SQUAT
JAD	BLADE	DOD	FED	CRUD	TWEED	FRIED	SQUAD
CAT	DATE	CLOT	GET	MUTT	WEET	KRAIT	SHAT
CAD	DADE	CLOD	GED	MUD	WEED	CRIED	SWAD
GNAT	DASE	POT	KET	SHUT	WHIT	PIGHT	WATT
NAD	DAZE	POD	KED	SHOULD	WHID	PIED	WAD
PAT	SOOT	PLOT	PET	SCUT	BRIT	SNITE	FEUTE
PAD	SUD	PLOD	PED	SCUD	BRID	SNIDE	FOOD
PLAT	CADE	SOT	STET	STUT	GRIT	TIGHT	HOOT
PLAID	CATE	SOD	STEAD	STUD	GROD	TIDE	HOOD
RAT	PATE	SKOT	THREAT	BLEAT	KIT	TRITE	MOOT
RAD	PAID	SKOD	THREAD	BLEED	KID	TRIED	MOOD
WAT	RATE	SQUAT	TRET	CHESE	QUIT	CROSE	FOOT
WAS	RAID	SQUAD	TREAD	CHEESE	QUID	CROZE	FOOD

Set B

				BARM	TAR	SAVLE	LA
				BALM	TA	SERVE	LAW
SHRAG	CARVE	SHRINE	MON	BARSE	HEARSE	UGH	DRAUGHT
SWAG	CALVE	SWINE	MUM	BASS	HUSS	HER	DROUTH
CHERT	CLART	SHRIVE	MONT	BARTH	SIR	FUM	THEE
CHAT	CLAUT	SWIVE	MENT	BATH	SO	FROM	THY
HAULSE	MARL	SOURCE	PURSE	CHAR	SEER	DUD	TIE
HOUSE	MALL	PSOAS	PUS	CHA	SEA	DID	TAILLE
HAULT	PARSE	FAULT	SHONG	DART	SHRIFT	GUN	KINE
HOUT	PASS	FOUGHT	SHUN	DOT	SWIFT	GON	KIN
GOLF	SCARP	GAULT	THIS	GAR	SHRILL	HUFF	RETCH
GOFF	SCAUP	GHAUT	THUS	GAW	SHILL	HAVE	REACH
ARSE	SMARM	SURE	AL	HAUGH	SHRINK	HUZZ	ILL
ASS	SMALM	SHIRR	ILE	HARK	SWINK	HAS	ILE
BARGH	SPAR	SPEARE	DEE	JAR			
BAFF	SPA	SPHERE	DIT	JAH			

Set C

AT	WADE	CUT	DID	CODE
BAT	KOD	NOULD	FID	LOTE
BRAT	GOD	PUD	GED	NODE
DRAD	QUAT	RUT	KIT	TOTO
LAD	NOD	SHOOT	CID	SHOAT
MAD	ROD	ESE	BIDE	BOOT
SCAD	SWAD	FEED	BRIDE	BROUD
BLOUSE	TOD	GLEET	GUIDE	LEUD
FADE	WAD	GREED	HIDE	HARD
GADE	BAWD	NEAT	SIDE	CARD
GRADE	RET	REIT	SICE	SAID
HADE	SAID	CEASE	SLIDE	BIDE
LATE	IDE	SWEDE	WIDE	BRIDE
MATE	BUD	IT	OAT	GUIDE
SPADE	FUD	BID	BODE	HIDE

Set D

DALT	HARRE	AR	°CAID	*MORE
AR	HARM	HER	°HOLD	°HOW
BARK	CARF	AYAH	*HAULM	*YOUR
GUAD	MAR	SOY	°HORSE	

Set E

AIT	AUGHT	GOT	GRETT
EIGHT	GUTTE	GHAIT	BRET
EYGHT			

DEBT	LET	GUT	BEAT
PETTE	LETT	GUTTE	BEET
HEAT	LEET	METE	CETE
HETE	LEAT	MEET	SEAT
		MEAT	
NIGHT	WRITE	GOAT	MODE
KNIGHT	RIGHT	GOTE	MOD
BOUGHT	ROOD		
BOTT	RUDE		

Set F

BAR	PAR	EARN
BARR	PARR	URN

Set G

CHAD	DOWD	BIRT	HERD
FORD	CORT	WARD	

Set H

SWATH	FART	SPOUT	SWEAT
SWATHE	FAD	SPALD	SHRED
GIRT	CURT	WORT	GIRT
GIRD	CURD	WORD	BIRD
CURT	CEAT	TIT	SORT
CURD	SURD	TEAT	SWORD

Set I

GHAUT
GALT

References

1. Dolby, J., and Resnikoff, H., "On the Structure of Written English Words," *Language,* Vol. 40, No. 2 (April-June, 1964).
2. *Webster's Third New International Dictionary of the English Language.* Springfield, Mass.: G. C. Merriam Co., 1961.
3. Kenyon, J. S., and Knott, T. A., A *Pronouncing Dictionary of American English.* Springfield, Mass.: G. C. Merriam Co., 1958.
4. *The American College Dictionary.* New York: Random House, 1962.
5. Jones, Daniel, *Everyman's English Pronouncing Dictionary.* 12th ed. New York: E. P. Dutton & Co., 1963.
6. *The Shorter Oxford English Dictionary on Historical Principles.* 3d ed., revised with addenda. Oxford: Clarendon Press, 1959.
7. Bhimani, B. V., and Mitchell, R. P., "Computable Relations between Orthographic and Phonetic Forms of English Monosyllables," unpublished manuscript available from the authors at Organization 52-40, Bldg. 201, Lockheed Palo Alto Research Laboratory, 3251 Hanover Street, Palo Alto, California.