WORD NETWORKS (PART 2)

A. ROSS ECKLER
Morristown, New Jersey

In the May 1973 issue of Word Ways, we presented the first part of a two-part article exploring the properties of word networks. A word network is formed out of words all having the same length; the words are joined by a link in the network if they differ only in one position, such as SHARE and SHORE, or BETTER and BATTER.

In the earlier article, we examined word networks consisting of 2-letter, 3-letter and 4-letter words drawn from the Merriam-Webster Pocket Dictionary. In this article, we conclude our study with a description of word networks for words from five to eight letters long.

Five-Letter Word Networks

The five-letter word network is the most difficult one to study. It is far too large to diagram, but on the other hand it is a tedious task to try and identify all the isolanos as was done in the four-letter case. However, one can employ statistical sampling methods to determine the size of the main network. The 58 words at the bottom of each column (of 48 words) in the Word Builder’s Handbook were carefully analyzed to see whether or not they could be joined to each other by any path using Pocket Dictionary words. After a considerable amount of effort, 33 of these were found to be connected; none of the remaining 25 were connected with the main network or with each other. Since there are a total of 2798 five-letter words, we estimate that the main network has a size of \(2798 \times \frac{33}{58} = 1590\) words, only slightly fewer than the four-letter main network. (For the statistically-trained reader, we point out that one can say with a 95 per cent probability of being correct that the true, but unknown, size of the main network lies in the interval between 1539 and 1641.)

Although the five-letter word network is too large to diagram, one can obtain some idea of its general structure. To begin with, words starting with a vowel play a relatively small role -- of the 33 sampled words, only three (AGAVE, EAGER and INTER) showed up. AGAVE (with AGAPE) is at the end of a relatively short vowel-chain: AGAVE-agate-abate-abase-abash-awash-SWASH. However, EAGER and INTER are contained in the most extensive vowel-chain connected to the main network (see next page). Note that this 25-word vowel chain joins the consonantal part of the main network in four places, and that all five initial vowels are presented.

We list below four of the groups, not almost immortal:

1) word
2) word
3) word
4) word

We list below the properties. The consonant, most that are fail other in the

--muc--

The only ot: longer; ign decay-deca: The side lo.
The first part of the main network has even more interesting properties. Letting C stand for a consonant, V for a vowel, and Y for itself, most of the consonantal network can be split up into four groups that are fairly strongly connected internally but weakly joined to each other in the order given by the following list:

1) words beginning with two consonants (CCVCV, CCVCC, CVCCV)
2) words with second letter a vowel, fourth letter a consonant (CVVVCC, CVVCV, CVVCY, CVCCY)
3) words with second letter a vowel, fifth letter Y (CVCCY, CVCCV, CVVCY)
4) words with second and fourth letters vowels (CVVCY)

We list below the only known essential connections between these groups. By essential connections, we mean those which anchor firmly in both groups, not filaments from one group that cross the line and terminate almost immediately.

```
CCVCC to CVVVCC: blast-boast
CCVCV to CVVVCC: place-peace, phase-pease, chase-cease, glide-guide, spite-suite, prise-poise
CVCCV to CVCCY: paste-pasty, haste-hasty, taste-tasty, mange-mangy, range-rangy
CVVCY to CVVCY: mouse-mousy, louse-lousy
CVCVY to CVCVC: decay-decal, honey-honer
```

Probably the most direct route from the second group to the third and fourth groups is the one from MOUSE to BOWER:

```
MOUSE-LOUSE
mousy--lousy
--mussy-mossy------bossy-bosky
mosey
money
honey-hooey
goner-honer
boner-borer
BOWER
```

The only other known route from the third group to the fourth is a bit longer; ignoring various short side paths, it runs BELLY-belay-delay-decay-decal-fecal-fetal-fatal-natal-naval-navel-novel-hovel-HOVER.
The word-supply in the Pocket Dictionary were a little smaller, the five-letter main network would very likely disintegrate into a number of large fragments. For example, the elimination of MOSEY and DECAL would cast the entire fourth group adrift.

What is the largest group of words isolated from the main network? This is quite difficult to determine; the following eighteen-word group is the largest one yet found:

- world
- wound
- wound-would
- poult
- mound
- mould
- mould-mould-mould-mound
- sound
- could
- count
- court
- found
- count-court
- round
- bound
- bound

The next largest isolated group may well be the eleven-word bigot-bight-eight-flight-light-might-night-right-sight-tight-wight.

The decreasing density of the five-letter network is well-illustrated by the fact that the word having maximum ambiguity, SHARE, connects with only fifteen words instead of twenty-four as was the case with WARE; SPARE, SCARE, SNARE, STARE, SHIRE, SHORE, SHADE, SHAKE, SHALE, SHAME, SHAPe, SHAVE, SHARD, SHARK and SHARP. Note that for the first time the maximum-ambiguity word does not have garbles in every position. Such words were dubbed ona10sis (the reverse of isolanos) by Dave Silverman in the May 1970 Word Ways. An example of a five-letter ona10sis is SHORE: CHORE, SCARE, SHARE, SHONE, SHORT. At the other extreme, sixteen of the twenty-five words sampled outside the main network were found to be isolanos; therefore, it is estimated that (16/58) = 27 per cent of all five-letter words are isolanos.

Garble groups play a steadily less important role. The largest known five-letter garble group, of shape (15211), has only ten words:

- share
- scare
- spare
- stare
- share
- shore
- score
- spore
- store
- shore

Murray Pearce has pointed out that the following set of words of shape (12241) lacks only two words for a sixteen word garble group:

- stove
- stave
- shove
- shave
- store
- stare
- shore
- share
- stoke
- stake
- shake
- stole
- stale
- shake

Note that this contains a garble group in which three letters are changed.
In the May 1972 Word Ways, Mary Youngquist presented a list of seventeen five-letter words with maximum resistance to ambiguity: ANGST, BATTER, LADYF, DRINK, ETHYL, FJORD, GIZMO, HELVE, ICTUS, LYNCH, OXBOY, PSALM, RUMBA, SPRIG, THEGN, UDDER, and WACKY. Of these words, only two -- DRINK and LYNCH -- are in the main network.

Six-Letter Word Networks

When one considers six-letter words, it is no longer accurate to speak of the main network. The 4056 words in the Pocket Dictionary are split into hundreds of independent fragments, the largest of which contains only 465 words -- 11.2 per cent of the total vocabulary. (It is quite likely that a few additional words have been overlooked, so 465 should be regarded as a lower bound on the group size.) Although this is far too large to conveniently diagram in this article, a brief description of its main characteristics follows. The great bulk of the words contained in it are of the form CVCCER, CVCCED and CVCCET, where C stands for a consonant and V, a vowel; the middle consonants are typically TT, NN, LL, GG, DD, RR, NG, CK, NK, RK, NC, ND, TH, SH, ST, LT, RT, RN or NT. About thirty words of the form CVVCER are also present. A group of 64 words of the form CVCCLE are joined to the CVCCER words via a single bridge: SETTLE, SETTEE, SETTER.

The second largest group is also composed of words of the form CVCCER, CVCCED and CVCCET, but the middle consonants are different: MM, PP, BB, MP or MB. It contains at least 101 words.

The third largest group, containing at least 66 words, is of the form CVCCLE with the middle consonants BB, MB, MP, RB, RG, RD or OD. The rather obvious group of CVCING words appears to be much smaller -- only 32 words are connected. It is possible that one or more unrecognized groups exceed this in size.

What six-letter word has maximum ambiguity? Surprisingly, BATTER is nearly as ambiguous as SHARE, with 13 words only one step away: HATTER, LATTER, MATTER, PATTER, TATTER, BETTER, BITTER, BUTTER, BANTER, BATHER, BATTED and BATTEN. Note that BATTER is not quite an isol osi, as it has no garble in the fifth position. It does not appear possible to find a six-letter isol osi using only Pocket Dictionary words; some near misses are:

HANGER: danger, hunger, -------, hanker, hangar, hanged
CANTER: banter, center, caster, canker, cantor, -------
BETTER: fetter, bitter, beater, -------, bettor, betted

If the word PASTER existed in the Pocket Dictionary, it would have garbles in every position: master, poster, patter, passer, pastor, pastel.

At the other extreme, a sample of 84 words revealed that 33 were isol anos; therefore, it is estimated that (33/84) = 39 per cent of all
six-letter words are isolanös.

What is the span of the largest word network? (It is rather unlikely that any of the spans of the smaller networks exceed this.) Although it is difficult to be certain, the following thirty-step path connecting SUDDEN to GIGGLY is probably the longest minimum-length path:


A path of similar length can be achieved by starting with BEDDED instead of SUDDEN. Although the six-letter network has slightly fewer words than the three-letter network did, the span is two-and-one-half times as large (30 compared to 12) -- evidence of the vastly different nature of the two networks.

It is possible to find fifteen six-letter words with maximum resistance to ambiguity: ASTHMA, BLAZON, CYSTIC, EMBRYO, FREEZE, GUFFAW, HICCUP, KNOBBY, LENGTH, MAddER, OBLONG, SCRUFF, TOWARD, UPHILL and WHILST.

If one has two six-letter words with maximum resistance to ambiguity, it is clear that it must take at least six steps to change one into the other by means of a word ladder -- one must travel along at least six lines in the network. Is it possible to find two such six-letter words which require no more than six steps to connect? This question can easily be answered in the affirmative for shorter words, but now it is not quite so trivial. In fact, only four such changes are known:

settle, settee, setter, better, batter, banket, banner
settle, settee, setter, better, batter, banter, banker
settle, settee, setter, better, batter, banker, banket
settle, settee, setter, better, batter, banter, banker, banket

Seven-Letter Word Networks

Not surprisingly, the fragmentation process continues with seven-letter networks. Although there are 4591 seven-letter words in the Pocket Dictionary, more than any other word length, the largest network has only 118 words -- about 2.6 per cent of the total. (The caveat given for six-letter words applies here as well.) A diagram of this word network is given on the next page. As all of the words but one (MATTINS) are of the form CVCCING or CVVCING, where C is a consonant and V, a vowel, the final three letters are omitted to save space.

Three other fair-sized word networks have also been noted: 47 words mostly of the form CCVCCER and CCVCCED; 35 words of the form CVVPFER and CVVPPED; and 24 words of the form CVPPING. Three of these four networks do not even collectively exhibit the on losers property; that is, one (or more) letters of all the words in each part
The seven-letter word with maximum ambiguity appears to be PETTING, which has eleven garbles: BETTING, GETTING, JETTING, LETTING, SETTING, WETTING, PATTING, POTTING, PUTTING, LONGING. Note that five of the seven letters are ungarbled. At the other extreme, a sample of 95 words revealed that 57 were isolanos; therefore, it is estimated that $(57/95) = 60$ per cent of all seven-letter words are isolanos.

As the size of the network shrinks, so does its span. The longest path is probably the one connecting PARTING and LONGING, with seventeen steps:


of these networks remain the same. However, the 47-word CCVCCER-CCVCCED network does produce garbles for every letter: platter-flatter, shatter-scatter, bluster-buster, planter-plaster, charter-charmer, chantey-chantry, and chatter-chattel.
It appears to be impossible to find two seven-letter Pocket Dictionary words with maximum resistance to ambiguity which can be connected in seven steps using Pocket Dictionary words. In fact, in the four networks described above, not one contains two seven-letter words with maximum resistance to ambiguity. It is conjectured that there exists no Pocket Webster network of seven-letter words with this property. Nevertheless, it is interesting to note that the 47-word CCVCCER-CCVCCED network almost accomplishes this feat:

chantry-chantey-chanter-chatter-shatter-sputter-
chantry-chantey-chanter-chatter-clatter-platter-plotter-

All words in both sequences have a T in the fifth position.

Eight-Letter Word Networks

We conclude this survey of word networks with the eight-letter version; it seems quite evident that for larger numbers of letters the networks will be quite small and uninteresting. (Nevertheless, one must go as high as sixteen-letter words in the Pocket Dictionary to find a pure set of isolano; fifteen-letter words have the networks INTERVENTIONISM-INTERVENTIONIST and SUPERINTENDENCE-SUPERINTENDENCY.)
The largest eight-letter word network is almost certainly the one which contains 28 words of the form CCVPPING, where C is a consonant and V is a vowel. Because only the first three letters change, it is possible to diagram this network in the style of the three-letter network (see preceding page).

The second largest network contains 14 words of the form CVGNESS and CVCCLESS, and the third largest network contains 12 words of the form CCVNNING and CCVNKING.

The word having maximum ambiguity in this network is clearly SLIPPING, with the seven garbles CLIPPING, FLIPPING, SHIPPING, SNIPPING, SKIPPING, SLAPPING and SLOPPING. Is it possible that an eight-letter word with greater ambiguity exists in some other network? At the opposite extreme, a sample of 91 words revealed that 74 were isolanos; therefore, it is estimated that (74/91) = 81 percent of all eight-letter words are isolanos.

The span of this network is ten: WRAPPING-trapping-tripping-dripping-dropping-cropping-chopping-shopping-slapping-SWAPPING.