

## THE POTENTIALITIES OF A MECHANICAL THESAURUS

### Summary

1. The logical effect which adopting the logical unit of the M.T. chunk, instead of the free word, has on the problem of compiling a dictionary.
2. Dictionary trees: an example of the tree of uses of the Italian chunk PIANT-.
3. As soon as the contexts of a second language are introduced, a dictionary tree becomes a lattice. The analogy between a multi-lingual dictionary lattice and an entry in a target-language Thesaurus.
4. Outline of a Mechanical Translation-programme using a Thesaurus.
5. Examples of trials made with a model-procedure for testing this: translations of ESSENZ-E, GERMOGL-I and SI PRESENT-A from the Cambridge Languages Unit's current pilot-project. The simplifications which the use of a Thesaurus makes in the research needed to achieve idiomatic Machine Translation.
6. Some preliminary remarks on the problem of coding a Thesaurus.

1. In M.T. literature, it is usually assumed that compiling an M.T. dictionary is, for the linguist, a matter of routine; that the main problem lies in providing sufficient computer-storage to accommodate it. Such judgments fail to take account either of the unpredictability of language, (Reifler, 1956), or of the profound change in the conception of a dictionary produced by the substitution of the M.T. chunk for the free word (Firth, B.B.C. Talk, 1955).

By chunk is meant here the smallest significant language-unit which (i) can exist in more than one context, and (ii) which, for practical purposes, it pays to insert as an entry by itself in an M.T. dictionary. Extensive linguistic data are of ten required to decide when it is, and when it is not, worth while to enter a language-unit by itself as a separate chunk. For instance, it has been found convenient to break up the Italian free word piantatore into the chunks PIANT-AT-ORE. It has not been found convenient to break up the Italian free word agronomi into chunks AGRO-NOM-I, but only into the chunks AGRONOM-I, since the addition of -NOM- to -AGRO- enables the distinction to be made between AGRO-, meaning "agriculture", and AGRO-, meaning "bitter",

Experience shows that the cutting-down of the number of entries, and the compensatory extension of the range of uses of each entry, caused by the substitution of chunks for free words, are together sufficient to call in question the current conception of a dictionary-article. In this paper we shall speak of current dictionary-articles, M.T. dictionary-entries, and Thesaurus items.

2. From the logical point of view, it can be shown that the range of uses of any chunk form a tree. Some paths of this tree are open to alternative analysis, but a considerable number of the paths, as of the points, can be

determined on objective criteria determined by the immediate context. For instance, the use of the Italian chunk PIANT- in the free word piantatoio is clearly different from its use in the free word piantatore. Moreover, the design of the tree can often be tested by its predictive value; for instance, in making the tree of the chunk FIBR-, a junction-point which had to be inserted to account for well-established uses was later found, when a larger dictionary was consulted, to be exactly fitted by the use of FIBR- in the free word fibroso, which had not appeared as an article in the smaller dictionary.

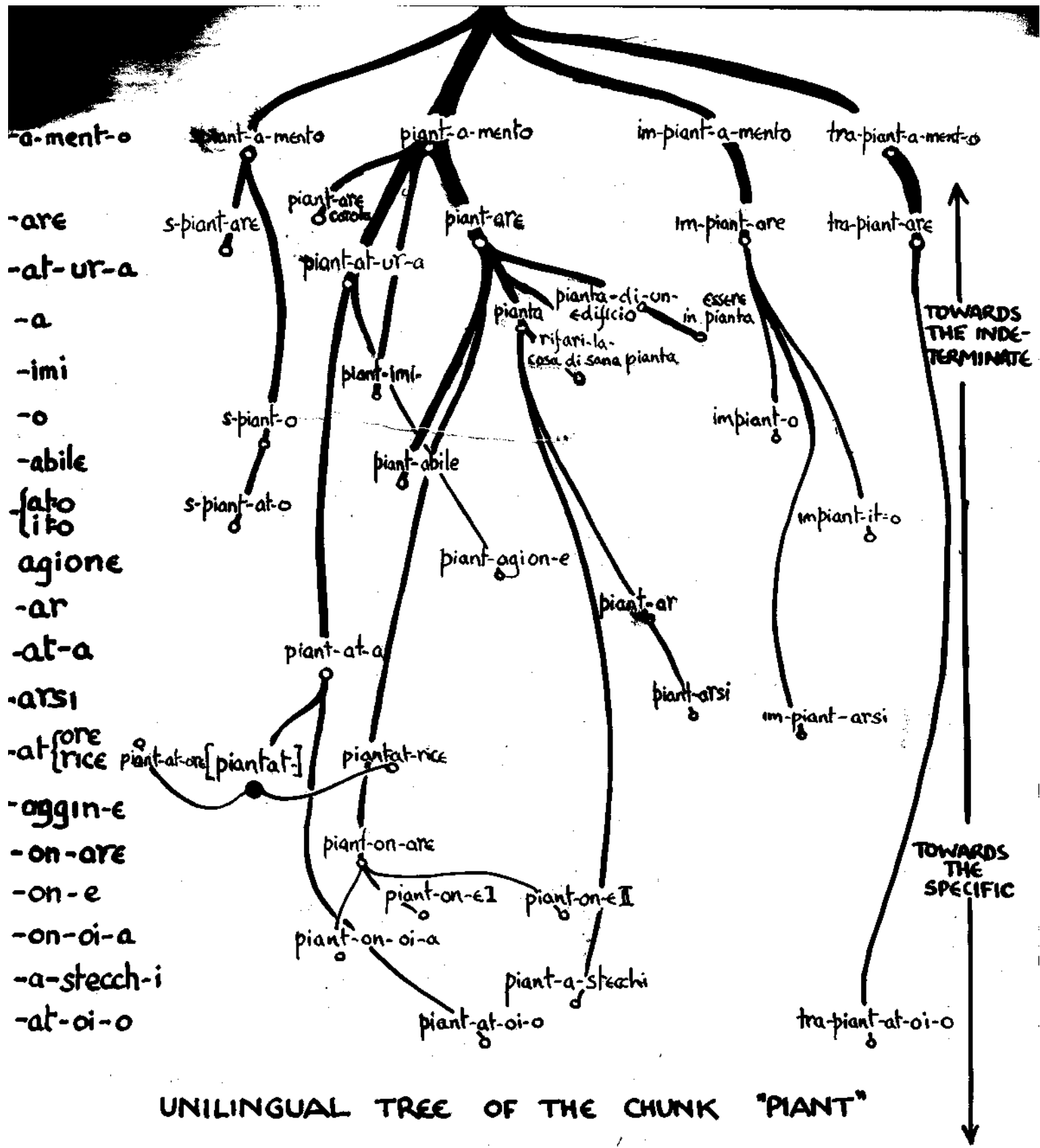
A unilingual tree of the Italian chunk PIANT-, compiled by M.M.Masterman and N.K.Willson, is given below. The specification of the dictionary-articles which the tree covers is taken from Lysle's Italian-English dictionary.

\*INSERT TEXT FIG.I.\*

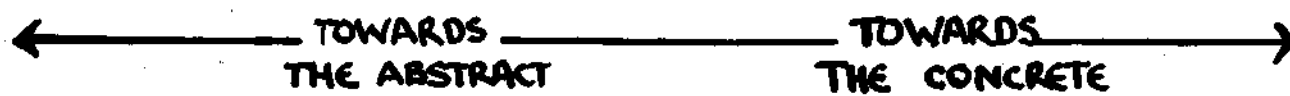
SUBSCRIPT TO TEXT FIG.I.\*  
\*PLEASE REDUCE\*

Dictionary-articles which contain the chunk PIANT-

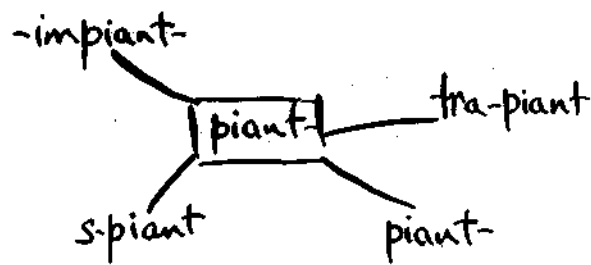
1. im-PLANT-ament-o, s.m., implantation, building, establishment.
2. im-PLANT-are, v.tr., to establish, to settle down to business, to found. impiantare una scrittura, to open an account.
3. im-PIANT-arsi v.rifl., to take one's stand.
4. im-PIANT-it-o s.m., floor, tiled place.
5. im-PIANT-o, s.m., establishing, setting up of a business.
6. PIANT-a, s.f., plant; tree: (arch:) plan, groundwork; sole (pianta dei piedi); lineage \*i.e. family tree\*: fig race: pianta esotica, exotic plant; pianta di un edificio, plan of a building; essere in pianta, to be on the list, rifare una cosa di sana pianta, to do a thing a second time.
7. PIANT-abil-e, adj., pertaining to a plantation.
8. PIANT-aggin-e, s.f., plantain, i.e. \*pasture-plant\*.
9. PIANT-agion-e, s.f., plantation, planting: piantagione di patate. potatofield.
10. PIANT-ament-o, s.m., planting, plantation.
11. PIANT-are, v.tr., to plant; to set; to stick; to drive in; to place; to forsake, to abandon, \*cf. French plager\*. piantare una bandiera, to set up a standard, to hoist a flag. piantare in asso, to leave a person in the lurch, piantare un pugnale nel petto, to stab with a dagger \*cf English, stick a dagger into him\*. piantare carote (fig) to make someone believe \*cf English, to plant a clue\* piantar le tende, to lodge, to dwell.
12. PIANT-arsi v.rifl., to fix oneself, to settle: to set up; piantarsi in un luogo, to settle down in one place.
13. PIANT-a-stecch-i, s.m., (calz.) punch, puncheon (arnese per piantar gli stecchi nelle suole).
14. PIANT-at-a, s.f., plantation; row of trees.
15. PIANT-at-o, part.pass.e.adj.planted, set up: fixed: ben piantato, well-built, well-set-up man.
16. PIANT-at-oi-o, m., (agr.) tool for planting \*dibbler\*.
17. PIANT-at-ore,-trice, s.m.f., planter.
18. PIANT-at-ur-a, s.f., plantation, planting.
19. PIANT-im-i, s.m., plur., many sorts of plantations.
- ( PIANTO , s.m., tears, weeping; lament; (fig) pain; regret.)
- ( PIANTO, part. pass., wept; lamented; deplored.)



UNILINGUAL TREE OF THE CHUNK "PIANT"



a three dimensional form of piant would go as follows



20. PIANT-on-ai-o, PIANT-on-ai-a, s.m.f., (agr.) nursery.
21. PIANT-on-are, v.tr., to watch over, \*to nurse, to guard\*: to plant cuttings.
22. PIANT-on-ie, s.m., (mil) sentry \*nurse, guard\*, (fig) watcher; (agr.) sucker scion, sapling. Essere di piantone, to sentine, \*to be on !! guard, to guard.\*
23. s-PIANT-a-ment-o, s.m., uprooting, transplanting.
24. s-PIANT-are, v.tr., to uproot, to transplant; to ruin.
25. s-PIANT-at-o, s.m., penniless person; (fam) someone who is dead broke, opp. stony broke.
26. s-PIANT-o, s.m., ruin, destruction. Mandare a spianto, to ruin.
27. tra-PIANT-a-ment-o, s.m., transplantation.
28. tra-PIANT-are, v.tr., to transplant
29. tra-PIANT-at-oi-o, transplanter (a tool).

\*.....\* translations added to the dictionary  
by M.M.M. and N.K.W.

When the contexts provided by translation into a second language are added to the above, the tree becomes very much more complicated. Inspection immediately shows, moreover, that the only criterion for differentiating many of the new points on the bi-lingual tree is the fact that, if, say, of two otherwise similar uses of PIANT-, English translations are given, different English words will be used in the two cases. For instance, once the English language is considered as well as the Italian, the use of PIANT- in the phrase *piantar le tende*, "to pitch a tent", must clearly be distinguished from its use in the phrase *piantare una bandiera* "to set up a standard". But to the man thinking wholly in Italian, this difference of use may not be perceptible: for him, one plants a tent on the ground and a standard in the air in exactly the same figurative sense of "plant"; all the more so, indeed, as *piantar le tende* means permanently to establish a tent (compare "Caesar then established his winter quarters") and is to be contrasted with *rizzare le tende*, which means to pitch a tent with the intention of taking it up again in a short time, - and this last differentiation of context is one which we have not got in English.

Such considerations raised doubts of the validity of such translation-points on bi-lingual dictionary-trees, which led to the re-analysis of bi-lingual dictionary-trees not as trees but as lattices. For translation-points on a dictionary-tree are not just points on a single path but junctions of two paths; as, indeed, the contexts of the uni-lingual tree might also be taken to be if such chunks as -UR- and -AGION- were taken as the points of origin of trees. Moreover, if it be granted that, even in simultaneous translation, translation is never actually made between more than two languages at once, a multi-lingual tree, as opposed to a bi-lingual tree, will also have this property that all its points will be translation-points, and it will therefore be a lattice. Moreover, it will not always be true that as the number of languages which are incorporated increases this lattice will become significantly more complex, because many of these translation-points will fall on one another.

3. In order to make clearer what is meant, a purely theoretical simplified schema of the multilingual lattice of the universal idea of physical grief in mourning is given below. In this each main chain within the lattice is meant to -represent the mourning metaphor used in a single language. This schema was included in a memorandum sent by M.M, Masterman to R.H.Richens, A.F.Parker-Rhodes and M.A.K.Halliday.

In his reply M.A.K.Halliday pointed out that a multi-lingual article, thus conceived, was closely analogous to an item in a single-language Thesaurus. In order to show the analogy, the Thesaurus-item Disappointment item 509 in Roget's Thesaurus, (Penguin Reference Books, R 7) is also shown below with its elements roughly analysed as points on a lattice:

\* INSERT TEXT-FIG II\*

20. Piant-on-ai-o, Piant-on-ai-a, s.m.f., (agr.) nursery.
21. Piant-on-are, v.tr., to watch over, \*to nurse, to guard\*: to plant cuttings.
22. Piant-on-e, s.m., (mil) sentry \*nurse, guard\*, (fig) watcher; (agr.) sucker scion, sapling. Essere di piantone, to sentine, \*to be on !! guard, to guard.\*
23. s-PIANT-a-ment-o, s.m., uprooting, transplanting.
24. s-PIANT-are, v.tr., to uproot, to transplant; to ruin.
- 25. s-PIANT-at-o, s.m., penniless person; (fam) someone who is dead broke, opp. stony broke.
26. s-PIANT-o, s.m., ruin, destruction. Mandare a spianto, to ruin.
27. tra-PIANT-a-ment-o, s.m., transplantation.
28. tra-PIANT-are, v.tr., to transplant.
29. tra-PIANT-at-oi-o, transplanter (a tool).

\* „„....\* translations added to the dictionary  
by M.M.M. and N.K.W.

When the contexts provided by translation into a second language are added to the above, the tree becomes very much more complicated. Inspection immediately shows, moreover, that the only criterion for differentiating many of the new points on the bi-lingual tree is the fact that, if, say, of two otherwise similar uses of PIANT-, English translations are given, different English words will be used in the two cases. For instance, once the English language is considered as well as the Italian, the use of PIANT- in the phrase *piantar le tende*, "to pitch a tent", must clearly be distinguished from its use in the phrase *piantare una bandiera* "to set up a standard". But to the man thinking wholly in Italian, this difference of use may not be perceptible: for him, one plants a tent on the ground and a standard in the air in exactly the same figurative sense of "plant"; all the more so, indeed, as *piantar le tende* means permanently to establish a tent (compare "Caesar then established his winter quarters") and is to be contrasted with *rizzare le tende*, which means to pitch a tent with the intention of taking it up again in a short time, - and this last differentiation of context is one which we have not got in English.

Such considerations raised doubts of the validity of such translation-points on bi-lingual dictionary-trees, which led to the re-analysis of bi-lingual dictionary-trees not as trees but as lattices. For translation-points on a dictionary-tree are not just points on a single path but junctions of two paths; as, indeed, the contexts of the uni-lingual tree might also be taken to be if such chunks as -UR- and -AGION- were taken as the points of origin of trees. Moreover, if it be granted that, even in simultaneous translation, translation is never actually made between more than two languages at once, a multi-lingual tree, as opposed to a bi-lingual tree, will also have this property that all its points will be translation-points, and it will therefore be a lattice. Moreover, it will not always be true that as the number of languages which are incorporated increases this lattice will become significantly more complex, because many of these translation-points will fall on one another.

3« In order to make clearer what is meant, a purely theoretical simplified schema of the multilingual lattice of the universal idea of physical grief in mourning is given below. In this each main chain within the lattice is meant to represent the mourning metaphor used in a single language. This schema was included in a memorandum sent by M.M.Masterman to R.H.Richens, A.F.Parker-Rhodes and M.A.K.Halliday.

In his reply M.A.K.Halliday pointed out that a multi-lingual article, thus conceived, was closely analogous to an item in a single-language Thesaurus. In order to show the analogy, the Thesaurus-item Disappointment item 509 in Roget's Thesaurus, (Penguin Reference Books, R 7) is also shown below with its elements roughly analysed as points on a lattice:

\* INSERT TEXT-FIG II\*

Discussion of this analogy led to the suggestion that a multilingual M.T. programme might be developed (given an imaginary computer of indefinitely expandable magnitude) in which the multi-lingual dictionary might be replaced by a target-language Thesaurus. The general design of such a programme is given in the next section, below. Work on an interlingua which might form the basis of a multi-lingual M.T. programme had already been begun by R.H.Richens, (R.H.Richens, 1952, 1956). Roget's Thesaurus had already been used in a programme to make the Manchester University Computer write love-letters (Christopher Strachey, 1953).

4. In this design the chunks of the input-text are passed through four successive processes of transformation. The first stage of each of these consists of matching the chunks, in turn, with some sort of dictionary; there are thus four dictionaries used in succession in the programme. These are 1. bi-lingual pidgin-dictionary: 2. lattice inventory: 3. Thesaurus cross-reference dictionary: 4. Thesaurus.

In order to exemplify this whole mechanical-translation process in concrete form, the following test-procedure has been devised. Translation trials might be undertaken which, if M.T. is to develop as a subject in its own right, will provide the controlled empirical material which we so much need.

In the procedure described below, the lattice-inventory and programme, which is by A.F.Parker-Rhodes, will in the near future actually go through a computer. The Thesaurus used was Roget's Thesaurus, (1953 edition), amended and amplified according to the procedures given below. The general design was by Masterman, and the pidgin passage-dictionary by Masterman and Halliday. The matches were made by means of alphabetically-stacked packs of written cards, each containing the entry for one chunk. In the case of two or more entries starting with the same chunk, the longer entry, in the case of each pair, was stacked first.

Since the method of matching with the lattice-inventory is more complicated, an appendix explaining the chunk-interpretation of lattice-theory as it is being used in the Cambridge Unit, and made with special reference to the Italian paragraph which is used to illustrate the Thesaurus test-procedure, is attached to this paper.

The procedure was developed as follows: A paragraph from an Italian botanical paper was chosen, and divided into chunks as given below:

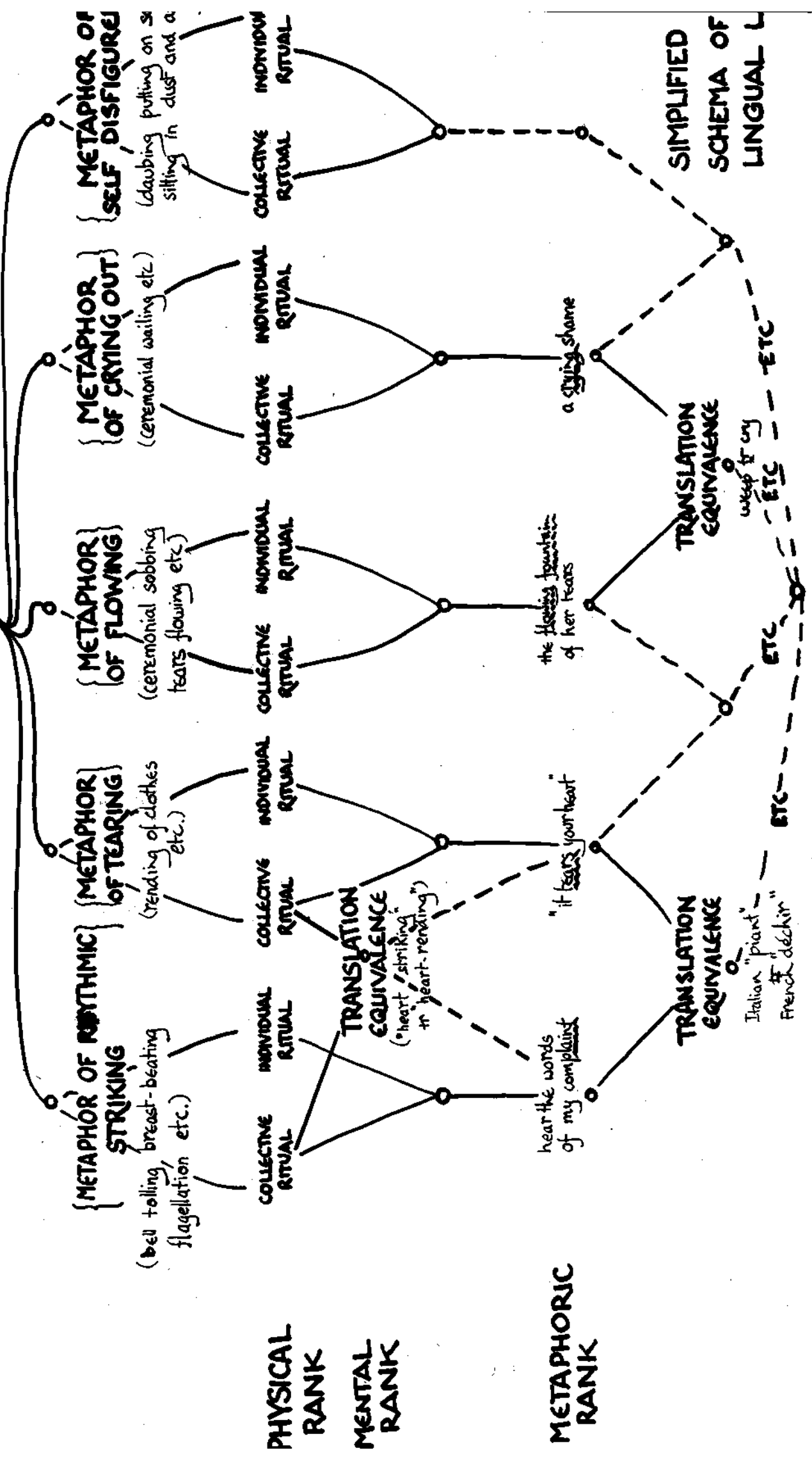
LA PRODUZ-ION-E DI VARIET-A DI PIANT-E PRIV-E  
 DI GEMM-E ASCELL-ARI , O PER+LE+MENO CON GERMOGL-I A  
 SVILUPP-O RIDOTT-O , INTERESS-A DA+TEMPO GENET-IST-I ED  
 AGRONOM-I, . TAL-E PROBLEM-A SI+PRESENT-A PARTICOLAR-MENT-E  
 INTERESS-ANT-E PER ALCUN-E ESSENZ-E . FOREST-AL-I E  
 FRUTT-IFER-I , PER LE PIANT-E DI FIBR-A , MA  
 SOPRATTUTTO PER IL TABACC-O. IN QUEST-A COLT-UR-A E  
 INFATTI IMPOSSIBIL-E MECCANIZZ-ARE L'-ASPORT-AZION-E DEI  
 GERMOGL-I ASCELL-ARI , NECESS-ARI-O D'+ALTRA+PARTE PER  
 OTTEN-ERE FOGLI-E DI MIGLIO-E QUALIT-A. .

(N.B. Entries of the form A+B+C..+N were entered as single chunks)

A simple Italian-English pidgin dictionary was then compiled covering the chunks of this paragraph. Specimen entries taken from this are given below. It will be noted that while the schema of this dictionary allows of one chunk having, if necessary, two Lattice Position Indicators, (L.P.Is),

# PHYSICAL GRIEF AFTER DEATH

## INTERLINGUAL SIGN



GENERAL  
INDETERMINATE  
IDEA

# DISAPPOINTMENT

ethically neutral concept

METAPHOR  
OF CROSSING

ABSTRACT  
NOUNS

CONCRETE  
CONCEPTS

LITERARY  
FORMS

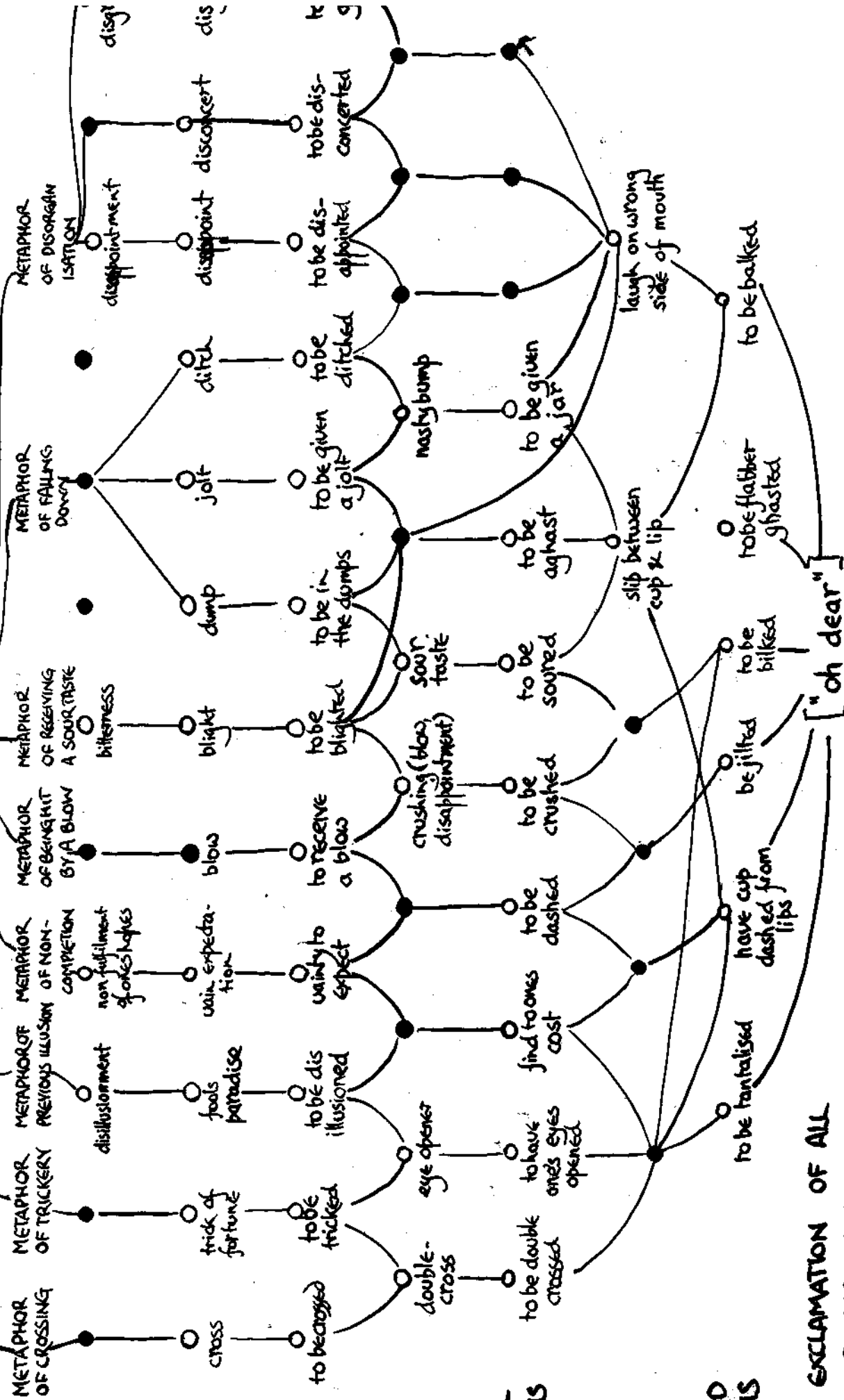
COLLOQUIAL  
NOUNS

COLLOQUIAL  
VERB FORMS

IDIOMS

GENERALISED  
VERB FORMS

GENERAL EXCLAMATION OF ALL  
KINDS OF DISAPPOINTMENT





though the chunks entered in this passage-dictionary have only one, it does not allow of any chunk having more than one pidgin translation. The numbers in the right-hand column govern a very simple first-approximative procedure for assigning singulars and plurals. The whole passage-dictionary was planned to give, as simply as possible, an output embodying only what the machine could immediately find out of the structure of Italian.

Sample Italian-English Pidgin-Dictionary Entries			
Italian	L.P.I.	1-> 0 routine	English
-A	28		o
-AL-	39		-Y
DA+TEMPO	28		FOR+SOME+TIME+PAST
FIBR-	30		FIBRE
I	26	0->1,1->1	THOSE-WHICH-ARE
GENET-	60		GENETIC-

When the chunks of this dictionary were matched with the chunks of the input, the following output was obtained:

Output I:                    top line;    singular/plural subroutine  
                                   second line;    output in chunks  
                                   decimal numbers;    L.P.I.s  
                                   initial set of subroutine (i.e. unmarked form) 1

```

{1->0
  26      29      54 28 35 {0->1,1->1
  THAT-ONE-WHICH-IS  PRODUCE-MENT-φ  OF  VARIETY-ω-S  OF PLANT-φ-S
  35      30 28      30 39 {set-back-to-1      56      28
  WITHOUT  BUD-ψ-S  AXIL-ARY  OR  AT+LEAST
  35      30 28      35      30      28 60 28 {set-back-to-1
  WITH  SPROUT-ψ-S  AT  DEVELOPMENT-χ-S  REDUCED-χ-S  ,
  29      28      28      60 30 28 56 62
  INTEREST-ω-S  FOR+SOME+TIME+PAST  GENETIC-IST-ψ-S  AND  AGRICULTURE-
  30 28 {set-back-to-1      31      {1->1,0  0->30  28 42      29      28
  IST-ψ-S  .  SUCH  16 PROBLEM-ω-S  SELF+PRESENT-ω-S
  60      35 29      39 28 35      26 28 30 28
  PARTICULAR-LY  INTEREST-ING-φ-S  FOR  SOME-φ-S  ESSENCE-φ-S
  30 39 28      56      30 39 28 {set-back-to-1      -      35      26
  FOREST-Y-ψ-S  AND  FRUIT-BEARING-ψ-S  ,  FOR THAT-WHICH-IS
  30 28 35      30 28 {set-back-to-1      -      56      28      35
  PLANT-φ-S  OF  FIBRE-ω-S  ,  BUT  ABOVE+ALL  FOR
  1->0,0->0 26      30 28 {set-back-to-1      31      35 {ignore-last-signal;
  THAT-ONE-WHICH-IS  TOBACCO-ω  .  IN  THIS  26 1->1, 0->0
  29      50 28 29-      28      60      28
  CULTIVATE-URE-ω  BE-α  IN+FACT  IMPOSSIBLE-φ
  29      -      26      29 28 35      30 28
  MECHANIZE-α  THAT-WHICH-IS  REMOVEMENT-φ  OF+THE  SPROUT-ψ
  30 39 {set-back-to-1      -      60 39 28      24      35
  AXIL-ARY  ,  NECESS-ARY-χ-S  ON+THE+OTHER+HAND  FOR
  29 48      30 28 35      28 28 30 28 {set-back-to-1
  OBTAIN-TO  LEAF-φ-S  OF  BETTER-φ-S  QUALITY-ωw-S  .
  31

```

The L.P.I.s of the chunks of this output were then picked up and inserted uniquely into lattices by means of the lattice-inventory and lattice-programme see appendix). These lattices give synthesis-routines for English which produce output II, below:

#### Output II

THE PRODUCE-MENT OF VARIETY-S OF PLANT-S WITHOUT AXIL-ARY BUD-S.  
OR AT+LEAST WITH SPROUT-S AT REDUCED DEVELOPMENT-S, (SING) INTEREST-(PRES)  
FOR+SOME+TIME+PAST GENETIC-IST-S AND AGRI CULTURE-IST-S (PLUR). SUCH  
PROBLEM-S (PLUR) SELF-PRESENT (PRES) PARTICULAR-LY INTEREST-ING FOR SOME  
FOREST-Y AND FRUIT-BEARING ESSENCE-S, FOR THE PLANT-S OF FIBRE-S, BUT  
ABOVE ALL FOR TOBACCO. IN THIS CULTIVATE-URE IT BE (PRES) IN+FACT  
IMPOSSIBLE TO MECHANIZE REMOVE-MENT OF+THE AXIL-ARY SPROUT,  
ON+THE+OTHER+HAND NECESSARY FOR TO OBTAIN LEAF-S OF BETTER QUALITY-S  
(PLUR).

It will be noticed that, in this output, the translation-procedure fails for non-grammatical reasons at a few easily identifiable points. (I am ignoring spelling-mistakes produced by the pidgin, such as PRODUCE-MENT for "production", as these could be picked up by cross-entries in the Thesaurus cross-reference dictionary). ESSENZ-E, in the original, is translated ESSENCE-S GERMOGL-I is translated SPROUTS; SI PRESENTA is translated SELF-PRESENT; and if ASCELL- had been given its vernacular meaning of ARMPIT-, the phrases ARMPIT-ARY BUD-S and ARMPIT-ARY SPROUT-S would have occurred in the translation\*

It was therefore decided further to examine these cases, by putting them through the Thesaurus cross-reference-dictionary and the Thesaurus.

Roget's Thesaurus cross-reference dictionary is arranged alphabetically. The entries in it form trees, but much simpler trees than those produced by normal dictionary-entries. Specimen entries from it are given below:

#### Specimen Entries from the cross-reference-dictionary of Roget's Thesaurus

<u>bud</u> 367	<u>plant</u>	<u>problem</u> 454, 461, 533,
beginning 66* 129*	place 184	<u>-atical</u> 475
germ 153	insert 300	
*ornament 847*	vegetable 367	
expand 194	agriculture 371	
graft 300	trick 545	
<u>-from</u> 154	tools 633	
<u>-dy</u> 711, 890	property 780	
	<u>-a battery</u> 716	
	<u>-oneself</u> 184	
	<u>-ation</u> 184, 371, 780	

It will be noticed that into the specimen entries given above, cross-references (between asterisks) have been inserted in the entries for bud and problem but not in the entry for plant. These insertions have been made to make the Thesaurus multi-lingua. They have not, however, been made ad hoc. If the Thesaurus dictionary procedure given here is to work for translation-trials, additions and emendations to the Thesaurus must be made only according to Thesaurus-principles; that is, according to one of the procedures given below:

---

\* In order to decide between AXIL- and ARMPIT-, as the translation for the pidgin-dictionary, a trial was made by rendering into pidgin the biblical story of Jeremiah the prophet, who was rescued from the pit by ropes which rested on the rags which he had put under his axils. This story remained comparatively comprehensible. This result could semantically, have been foreseen, since an armpit is an instance of an axil, as is also the crutch of the legs - the only other place Jeremiah could have put his rags, whereas the idea of an axil cannot, inductively, be reached from that of an armpit.

Procedure for amplifying a translation-Thesaurus

Each chunk in any pidgin-dictionary must successfully match with an entry in the cross-reference-dictionary: e.g, PLANT-, plant.

Each main meaning of the corresponding source-language entry in the pidgin; dictionary, must be compared (not matched) with the sub-headings of the cross-reference entry. If the comparison is unsatisfactory in that there is reason to suspect that the cross-reference-spread is too narrow/ (i.e. that the cross-reference-tree has not enough main branches) then one of the two emendation-procedures given below must be adopted.

(i) without making an addition to the cross-reference entry, bring down the actual Thesaurus-items which are referred to in the entry and search for the missing meanings. If they are found, no addition to the cross-reference entry need be made.

example The Italian bi-lingual dictionary tree of PIANT- (actually a lattice) has a branch with the main meaning design. This branch has derived meanings groundwork, plan, blue-print, installation; list; scheme, invention, pretext, lie. In the cross-reference entry "plant as design" does not occur, "plant as trick, 545", however, does; and the Thesaurus item 545, Deception, gives, either directly or by sub-reference, lie, pretext, invention, and blue-print. Scheme, design and plan, can also readily be reached from this item if (under emendation-procedure (ii). below) an addition is made to item 545, row 3, so that this row now reads:

item 545, row 3: trick, cheat, wile, ruse, blind, feint, plan, catch, chicane, juggle, reach, hocus; thimble-rig, card-sharping, artful dodge, machination, swindle, hoax, hanky-panky; tricks upon travellers; confidence trick; stratagem, &c. 702; \*scheme, &c, 626; \* theft, etc., 791.

That the new asterisked element is a legitimate addition to the Thesaurus can be confirmed by consulting item 702, Cunning, where schemer occurs, and where there is a reference back to 545.

List could legitimately be inserted into 626 as follows:

item 626, row 4: \*list\*, programme, \*&c, 86\*; forcast, play-bill, prospectus, scenario,....

This addition can be checked by looking up 86, List, which already contains programme.

List • should also be inserted into 626 in row 11, so that this now reads:

item 626, row 11: cast, recast, systematise, organise; arrange, \*list\*, &c, 60, digest, mature.

This addition can be confirmed by consulting item 60, Arrangement, which already contains list.

Finally, under list, in the cross-reference dictionary, a subheading must now be added "list as plan, 626", so that the total entry now reads:

list  
as catalogue 86  
\*as plan 626\*  
as strip 205  
as leaning 217,  
etc, etc.

Of the remaining meanings of PIANT-, routes to groundwork and installation can only be constructed, if at all, by more intervening steps, since the items 25, Support and 185, Location, where they occur, do not appear in the dictionary cross-reference entries of any of the others.

Thus there is no incentive to add "plant as design" to the cross-reference-entry of plant, which would be done under procedure (ii) since the entry "plant as trick" already leads to all the items which could thus be reached.

(ii) Under this procedure an addition is made to the actual cross-reference dictionary-entry of the chunk in question.

Example: The bi-lingual dictionary-tree (actually a lattice) schematising the uses of GEMM- contains a branch of which the main English meaning is "gem". The Thesaurus cross-reference dictionary-entry BUD includes no cross-reference which leads to any item containing gem or jewel. If, however, the cross-reference, "bud as ornament 847" is added between "bud as germ" and "bud as expand", (see above), the required connection is made, since item 847, Ornament, reads as follows:

item 847, row 7: tassel, knot; epaulet, aigulet, frog; star, rosette, bow; feather, plume, aigrette.  
row 8: jewel, jewellery; bijouterie; diadem, tiara; pendant, trinket, locket, necklace, armilla, bracelet, bangle, armband, anklet, ear-ring, nose-ring, chain, chatelaine, brooch.  
row 9: gem, precious stone; diamond, emerald; onyx, plasma; opal, sapphire, ruby; amethyst, pearl...

We now have the required connection from entry to item. In order to be able to get back from item to entry, however, one of the given rows of 847 must be extended so as to include bud. The suggested extension is as follows:

item 847, row 7 (contd:) ...feather, plume, aigrette; \*bump, button, nipple, nodule, bud.

The justification for this extension, of course, has got to be that some, at least, of this chain of metaphoric uses exist in English. Bump can be taken as colloquial: ("that is a very ornamental bump you have upon your forehead."). Ornamental buttons are dressmaking stock in trade; this element should be already in the item. Nipple has a definite, though rare, use as a nipple-shaped beautiful object, ("The crests or nipples of the hill line are crowned with the domes of the mosques, "wrote Cory in 1873: Oxford Dictionary). Nodule has an even rarer one, meaning "something like a knot". Finally, bud, meaning "ornament", does exist, but only poetically and archaically. Thus we get "Their breasts they embuske on high and their round Roseate buds immodestly lay forth," (Nashe, 1613). And Emerson, in his poems, wrote much later of "the bud-crowned spring".

Thus we get the curious situation that the use of an extended train of meanings for ornament, all of which have become cliches in Italian, is still an act of poetic originality in English and American.

Nevertheless, the train of uses exists, and the addition to the Thesaurus item is therefore justified.

These methods of emending and amplifying Roget's Thesaurus have been exemplified in detail, because, in view of the surprisingly good outputs which follow, it might be thought that the Thesaurus-routes used had been manipulated to suit the Italian paragraph. This is not so; every suggested new connection has been checked and justified, and all relevant asterisked emendations used to reach the outputs are given below. The suspicion of manipulation represents a direction opposite to that in which the research has gone, for, in actual fact, the more the experience which is gained of using this Thesaurus, the less the emendations which are made. It is a sound presumption that, with few exceptions, all possible chains of meanings are somewhere in Roget's Thesaurus if they can be found. A minimum number of trials, moreover, begets a strong conviction that Thesaurus searching and matching would best be done automatically from the earliest possible date; they are no work for a mere human being. In other words, if the thesaurus-technique proves, on trials, to have definitive M.T. significance, it will also prove to be the frontier-point where the M.T.worker, in this new kind of calculation, hands over to the machine; where results, uncalculated in advance by the programmer, are produced by the programme. It may also

(that is, if it establishes itself as having translation value), be the point of departure for a new exploration of the analogy between the human cortex and a computer; for this feels like a model of what we do when we ourselves translate.

5. Work done on the Italian paragraph has provided the following examples of translations produced by the Thesaurus procedure.

case I: ESSENCE-S.

If the chunks FOREST AND FRUIT-BEARING ESSENCE-S, - that is, all the chunks in the invertor-lattice 56, 60, 60, in which they occur (see appendix) are matched with the entries in the thesaurus cross-reference dictionary, the following output is obtained:

Output III:

forest \*57\* 367, 890 and 37, 38 fruit result 164  
 produce 161  
 food 298  
 profit 775  
 forbidden- 615  
 reap the -s, 973  
 -tree, 367  
 fruitful 168  
 fruition 101  
 fruitless 169. 645, 732

<u>bearing</u>	<u>essence</u> 5, 398
relation 9	essential
support 215	intrinsic 5
direction 278	*meaning 516*
meaning 516	great 31
demeanour 692	required 630
-rein 752	important 642
*fruit- 168, 637, 367*	essentially 3, 5
*child- 161*	essential stuff 5

Upon this output the thesaurus-operations are performed with the aid of restrictive and permissive rules, given as they occur, and the object of which will be evident. If the machine could be programmed to know that ESSENCE, and not FOREST-, or FRUIT-BEARING, is the word that needs to be retranslated, the right output, namely "example", would be obtained, because the machine could then be instructed to suspend any restrictive rule which is designed to prevent a chunk already rightly translated in Output II from being replaced by a string of synonyms. Such a rule would have to run, "In the case of the chunk to be retranslated, reject output given by Rule X, and replace by output normally rejected by rule X. We will call this rule Post-Editing Rule I, to show that, in this thesaurus-procedure, it cannot be automatised.

Operation 1.1 Pick out all numbers which occur more than once in Output III. Let these be called ring numbers.

result 1.1

<u>ring number</u>	<u>Thesaurus item</u>	<u>sources of ring number</u>
367	Vegetable	forest, fruit
161	Production	fruit, fruit, bearing
168	Productiveness	fruit, bearing
516	Meaning	bearing, essence
5	Intrinsicity	essence, essence, essence

It is worth remarking, as an incidental fact, that "The Intrinsic Meaning of the Productiveness of Vegetable Production" could stand as a sub-title, of a sort, for the whole paper.

Operation 1.2 Reorder ring numbers in order of descending frequency of occurrence. In the case of two ring numbers which occur with equal frequency, put first those which ring together most chunks. If order is then still undecided in any case, take input order.

result 1.2 5, 367, 161, 168, 516

Operation 2.1 Compare for common elements, in twos, the Thesaurus items bearing the ring numbers in the comparisons which are permitted by the lattice-relations of the chunks which are being put through the procedure, (in this case those of the inverter-lattice 60, 56, 60). In the case of any two chunks, A and B, call this comparison  $A \cap B$ ,

- Order of comparisons:
- (i)  $A \hat{=} A$  (e.g. fruit  $\hat{=}$  fruit)  
N.B.. When this lattice-relation yields  $\underline{a \cap a}$ ,  $\underline{a}$  being not a chunk but a ring number, take output which is identical with original chunk.  
(example  $\underline{161 \cap 161}$  )
  - (ii) A covers B
  - (iii)  $A \geq B$

The output produced by the comparison, subject to the restrictive and permissive rules given below, is to be taken as synonymous with the chunk A in the form  $A \geq A$ , and with the chunk B in the case where A covers B or  $A \geq B$ .

Since the inverter-lattice-elements 60, 56, 60 are formed from 2-element-chains 30, 39 (see Appendix), the following comparisons are permitted in this case.

<u>lattice-relation</u>	<u>chunk-comparison</u>	<u>ring number-comparison</u>
$A \geq A$		
$\equiv A \cap A = A$	FRUIT $\cap$ FRUIT	367 $\cap$ 161
	FRUIT $\cap$ FRUIT	161 $\cap$ 168
	FRUIT $\cap$ FRUIT.	367 $\cap$ 168
	BEARING $\cap$ BEARING	161 $\cap$ 168
	BEARING $\cap$ BEARING	161 $\cap$ 516
	BEARING $\cap$ BEARING	168 $\cap$ 516
	ESSENCE $\cap$ ESSENCE	5 $\cap$ 516
A covers B	FOREST $\cap$ -Y	No comparison, as -Y has no entry
	FRUIT $\cap$ BEARING	161 $\cap$ 168
	FRUIT $\cap$ BEARING	161 $\cap$ 367
	FRUIT $\cap$ BEARING	161 $\cap$ 516
	FRUIT $\cap$ BEARING	168 $\cap$ 367
	FRUIT $\cap$ BEARING	168 $\cap$ 516
	FRUIT $\cap$ BEARING	367 $\cap$ 516
$A \geq B$	FOREST $\cap$ ESSENCE	367 $\cap$ 5
$\equiv A \cap B = B$	FOREST $\cap$ ESSENCE	367 $\cap$ 516
$(A \cap B) \cap C$	FRUIT-BEARING $\cap$ ESSENCE	161 $\cap$ 5
$\equiv (A \cap C) \cap (B \cap C)$	FRUIT-BEARING $\cap$ ESSENCE	168 $\cap$ 5
	FRUIT-BEARING $\cap$ ESSENCE	161 $\cap$ 516
	FRUIT-BEARING $\cap$ ESSENCE	168 $\cap$ 516
	FRUIT-BEARING $\cap$ ESSENCE	367 $\cap$ 5
	FRUIT-BEARING $\cap$ ESSENCE	367 $\cap$ 516

N.B. The comparison FOREST  $\cap$  FRUIT is prohibited, since these chunks are incomparable in the lattice. But no new comparison would result from allowing this, since all possible combinations of the five numbers already occur.

Operation 2.2 List common elements given by Thesaurus-item comparisons.

<u>ring numbers</u>	<u>thesaurus-items</u>	<u>outputs</u>
5 $\cap$ 161	Intrinsicity $\cap$ Production	flower; &c 22
<u>New Comparisons Generated:</u>		
5 $\cap$ 22	Intrinsicity $\cap$	example, specimen
161 $\cap$ 22	Prototype	
	Production $\cap$ Prototype	pattern, prototype
5 $\cap$ 168	Intrinsicity $\cap$ Productiveness	NO OUTPUT
5 $\cap$ 367	Intrinsicity $\cap$ Vegetable	flower
5 $\cap$ 516	Intrinsicity $\cap$ Meaning	essence, example, meaning, &c 22
<u>New Comparisons generated;</u>		
5 $\cap$ 22		SEE ABOVE
516 $\cap$ 22	Prototype $\cap$ Meaning	prototype, example
161 $\cap$ 168	Production $\cap$ Productiveness	propagation, fertilisation, fructify, produce 161 - 168, 168-461
161 $\cap$ 367	Production $\cap$ Vegetable	growth, flower
161 $\cap$ 516	Production $\cap$ Meaning	prototype &c 22
<u>New Comparisons Generated:</u>		
161 $\cap$ 22		SEE ABOVE
516 $\cap$ 22		SEE ABOVE
168 $\cap$ 367	Productiveness $\cap$ Vegetable	NO OUTPUT
168 $\cap$ 516	Productiveness $\cap$ Meaning	NO OUTPUT
367 $\cap$ 516	Vegetable $\cap$ Meaning	NO OUTPUT

Operation 3.1. Produce synonyms for the passage required by applying outputs given under 2.2. to comparisons permitted under 2.1.

synonym-outputs

- for FRUIT
- (i) growth, flower
  - (ii) propagation, fertilisation, fructify, produce
- N.B. since cross-references both from 161 to 168 and from 168 to 161 lead to permitted comparisons 161  $\cap$  161 and 168  $\cap$  168, apply 2.1., i, and substitute FRUIT

for BEARING	AS ABOVE.'i.e., under 2.1.i, substitute BEARING
for ESSENCE	essence, example, meaning, &c 22 example, specimen prototype, example
for FRUIT-BEARING	AS ABOVE, i.e. under 2.1., i, substitute FRUIT-BEARING
for FOREST ESSENCE	flower
for FRUIT-BEARING ESSENCE	flower, &c 22, example, specimen, pattern, prototype, prototype, & 22, example, specimen, prototype, example, flower.

So far, we have used no restrictive or permissive rule except 2.1., i. If we make use of the following additional rules, to distinguish between outputs, we get the following final result:

- Restrictive Rules
- (i) 2.1., i (as above)
  - (ii) If a chunk of output II generates no ring number in the thesaurus, and thus generates also no comparison, replace it by itself in Output IV  
By this rule, FOREST is reinserted as FOREST
  - (iii) If rule 2.1., i operates, reject all other output.  
  
By this rule, FRUIT remains FRUIT, -BEARING remains -BEARING, and FRUIT-BEARING remains FRUIT-BEARING
  - (iv) When selecting final output, take longest output first, i.e. if there is a synonym-output for FRUIT-BEARING ESSENCE, select it in preference to a synonym for FRUIT-BEARING. (This is analogous to the pidgin-dictionary matching rule, given earlier).  
  
By using these, we remove all but the final synonyms

Output IV:

for FOREST ESSENCE	forest flower
for FRUIT-BEARING ESSENCE	fruit-bearing example, (3 occurrences), flower (2 occurrences), prototype (2 occurrences), specimen, (2), pattern, (1 occurrence).

N.B. In this output, alternatives have been reordered in order of occurrence, and the output &c 22 deleted.

Asterisked entries in Thesaurus: In item 5:

item 5, row 1: ...essence, essentialness,...essential part,...gist,pith,  
core, kernel, marrow,... important part, &c, 642, \*meaning,  
&c, 516\*

row 2: principle, nature, constitution, character, type, quality;  
\*token, example, instance, specimen &c 22;\*

item 161,row 4: authorship, publication, works, opus; \*result, answer,  
calculation; arrangement, pattern, prototype, &c 22;  
product, treatment\*

In the case of ESSENCE, the full Thesaurus test-procedure has been given. In the other oases taken from the Italian paragraph, which follow, only the results of the successive operations are shown.



Case II: SELF-PRESENT

1 .1. ring numbers

such 17  
problem 454

self 13, 79, 451, 486, 565, 604, 717, 836, 861, 864, 879, 880,  
942, 943, 950, 953, 990

present 118, 151, 186, 457, 505, 763, 861, 894

interest 454, 455, 457, 780

(1.2. re-ordering of these in descending order of frequency of occurrence)

2.1. permitted comparisons output  
(153 comparisons wait for future computer)

SELF  $\cap$   
SELF

PRESENT $\cap$	118 $\cap$ 151	Eventuality $\cap$ Present Time	NO OUTPUT
PRESENT	118 $\cap$ 186	Presence $\cap$ Present Time	present
	118 $\cap$ 457	Present Time $\cap$ Attention	NO OUTPUT
	118 $\cap$ 763	Present Time $\cap$ Courtesy	present
	118 $\cap$ 894	Present Time $\cap$ Offer	NO OUTPUT
	151 $\cap$ 186	Eventuality $\cap$ Presence	NO OUTPUT
	151 $\cap$ 457	Eventuality $\cap$ Attention	concern
	151 $\cap$ 763	Eventuality $\cap$ Offer	NO OUTPUT
	151 $\cap$ 894	Eventuality $\cap$ Courtesy	NO OUTPUT
	186 $\cap$ 457	Presence $\cap$ Attention	NO OUTPUT
	186 $\cap$ 763	Presence $\cap$ Offer	NO OUTPUT
	186 $\cap$ 894	Presence $\cap$ Courtesy	NO OUTPUT
	457 $\cap$ 763	Attention $\cap$ Offer	NO OUTPUT
	457 $\cap$ 894	Attention $\cap$ Courtesy	attentive
	763 $\cap$ 894	Offer $\cap$ Courtesy	NO OUTPUT
PARTICULAR $\cap$	79 $\cap$ 151	Speciality $\cap$ Eventuality	NO OUTPUT
PARTICULAR	79 $\cap$ 594	Description $\cap$ Speciality	particularise, specify
	79 $\cap$ 700	Speciality $\cap$ Property	personal
	51 $\cap$ 594	Eventuality $\cap$ Description	NO OUTPUT
	151 $\cap$ 780	Eventuality $\cap$ Property	business
	594 $\cap$ 780	Description $\cap$ Property	NO OUTPUT
INTEREST $\cap$	454 $\cap$ 455	Topic $\cap$ Curiosity	interest, &c, 461
INTEREST	454 $\cap$ 457	Topic $\cap$ Attention	&c. 451
	454 $\cap$ 780	Topic $\cap$ Property	interest, business
	455 $\cap$ 457	Curiosity $\cap$ Attention	interest, attentive
	455 $\cap$ 457	Curiosity $\cap$ Property	NO OUTPUT
	457 $\cap$ 780	Attention $\cap$ Property	NO OUTPUT

By these comparisons, two new ring numbers are generated, 451, 461. These cause the ring numbers for problem now to be 454, 451, 461, and the ring numbers for interest now to be 451, 454, 455, 457, 461, 780. These additions permit the following additional comparisons of the form  $A \geq A$ .

PROBLEM $\cap$	454 $\cap$ 451	THOUGHT $\cap$	&c 461
PROBLEM		TOPIC	
	451 $\cap$ 461	THOUGHT $\cap$	study, discuss, consider
		INQUIRY	
	454 $\cap$ 461	TOPIC $\cap$	&c 451; question, problem
		INQUIRY	

INTEREST $\cap$	451 $\cap$ 454	THOUGHT $\cap$ TOPIC	&c 461
INTEREST	451 $\cap$ 455	THOUGHT $\cap$ CURIOSITY	&c 457
	451 $\cap$ 457	THOUGHT $\cap$ ATTENTION	thought, reflection, consideration, inter- est, close study, occupy the mind, strike one as, &c 458
	451 $\cap$ 461	THOUGHT $\cap$ INQUIRY	study, discuss, consider
	451 $\cap$ 780	THOUGHT $\cap$ PROPERTY	NO OUTPUT
	454 $\cap$ 455	TOPIC $\cap$ CURIOSITY	NO OUTPUT
	454 $\cap$ 457	TOPIC $\cap$ ATTENTION	interest, &c 451
	454 $\cap$ 461	TOPIC $\cap$ INQUIRY	&c 451, question, problem
	454 $\cap$ 780	TOPIC $\cap$ PROPERTY	business
	455 $\cap$ 457	CURIOSITY $\cap$ ATTENTION	interest, attentive
	455 $\cap$ 461	CURIOSITY $\cap$ INQUIRY	prying, what's the matter?
	455 $\cap$ 780	CURIOSITY $\cap$ PROPERTY	NO OUTPUT
	457 $\cap$ 461	ATTENTION $\cap$ INQUIRY	&c 451
	457 $\cap$ 780	ATTENTION $\cap$ PROPERTY	NO OUTPUT
	461 $\cap$	INQUIRY $\cap$ PROPERTY	NO OUTPUT

At this point the detailed procedure was broken off, since it was already clear that the output of greatest frequency, among the synonyms given for INTEREST would be "thought, reflection, consideration, interest, close study, occupy the mind, strike one as, &c 458," namely the output of 451  $\cap$  457. For the additional newly generated ring number, 458, Inattention, yields only &c 457 as output when compared with any of the others; and this output is already also given by 451  $\cap$  455. Three other outputs already also include 451. Thus if the work of comparison is continued, the combination 451  $\cap$  457 will increasingly recur.

It is clear that the synonyms required for idiomatic translation of SELF-PRESENT namely "strike one as, occupy the mind", will occur in the wrong position, namely as synonyms for INTEREST. Nor can this error be corrected from the lattice-program. For this, as given, allows only the comparisons SUCH  $\cap$  PROBLEM and PARTICULAR  $\cap$  INTEREST, neither of which will improve the synonym-output for SELF-PRESENT. The only lattice-relations which will produce the required connection are those given by the extended lattice consisting of the whole sentence, and only this after the dualling operation, (see appendix) has already been performed. This operation, by reversing the meets and joins of the lattice, allows SELF-PRESENT  $\cap$  PROBLEM to occur as B-element of a 2-element chain of which PARTICULAR  $\cap$  INTEREST occurs as A-element and thus allows 2.1., v to operate. But this intersentential lattice-programme does not exist as yet.

The final output, therefore, of this application of the procedure, is as follows:

2,2	for PROBLEM	study, discuss, consider; question, problem
	for PRESENT	present, concern, attentive
	for PARTICULAR	particularise, specify, personal, business
	for INTEREST	thought, reflection, consideration, interest, close study, occupy the mind, strike one as; study, discuss, consider; question, problem; business; attentive; prying, what's the matter?
	for PARTICULAR	application, hobby, particularity, applica-
	$\cap$ INTEREST	tion, indicate, prove, occur, find, affair, run over, specification.

Of these last, the output of 151  $\cap$  451, Eventuality  $\cap$  Thought, prove, occur, find, is of interest, as it would be given under 2.1., v, above, since 151 is a ring number also in PRESENT.

case III: SPROUT

1.1. ring numbers

with 52

sprout 35, 154, 194

reduce 144, 160

development 35, 144, 154, 194

1.2. ring numbers in order of frequency of occurrence; 35, 144, 154,  
194, 52, 160

2.1. permitted comparisons

SPROUT $\cap$ SPROUT	35 $\cap$ 154	Increase $\cap$ Effect	production, development, grow, sprout, shoot
	35 $\cap$ 194	Increase $\cap$ Expansion	increase, enlargement, augmentation, extension, growth, development, spread, swell, shoot, sprout
	154 $\cap$ 194	Effect $\cap$ Expansion	growth, development, sprout, shoot.
REDUCE $\cap$ REDUCE	144 $\cap$ 160	Conversion $\cap$ Weakness	reduce
DEVELOPMENT $\cap$ DEVELOPMENT	35 $\cap$ 144	Increase $\cap$ Conversion	growth, development, grow
	35 $\cap$ 154	Increase $\cap$ Effect	SEE ABOVE
	35 $\cap$ 194	Increase $\cap$ Expansion	SEE ABOVE
	144 $\cap$ 154	Conversion $\cap$ Effect	grow
	144 $\cap$ 194	Conversion $\cap$ Expansion	development, growth, grow
	154 $\cap$ 194	Effect $\cap$ Expansion	SEE ABOVE
REDUCE $\cap$ DEVELOPMENT	SEE SPROUT $\cap$ REDUCE $\wedge$ DEVELOPMENT		
SPROUT $\cap$ REDUCE $\wedge$ DEVELOPMENT	35 $\cap$ 144	Increase $\cap$ Conversion	growth, development, grow
	35 $\cap$ 154	Increase $\cap$ Effect	production, development, grow, sprout, shoot
	35 $\cap$ 160	Weakness $\cap$ Increase	shoot
	35 $\cap$ 194	Increase $\cap$ Expansion	increase, enlargement, augmentation, extension, growth, development, spread, swell, shoot, sprout
	144 $\cap$ 154	Conversion $\cap$ Effect	grow
	144 $\cap$ 160	Conversion $\cap$ Weakness	reduce
	144 $\cap$ 194	Conversion $\cap$ Expansion	development, growth, grow
	154 $\cap$ 160	Effect $\cap$ Weakness	bud, shoot
	154 $\cap$ 194	Effect $\cap$ Expansion	growth, development, sprout, shoot

2.2. synonyms for SPROUTS, in SPROUTS AT REDUCED DEVELOPMENT:

development (5 occurrences), shoot (5 occurrences) growth,  
(4 occurrences) sprout (3 occurrences) production, bud,  
reduce, spread (1 occurrence).

Asterisked Entries in Thesaurus, for cases II and III:

cross-references:

interest

concern 9  
\*occupation\*  
curiosity 455  
etc.

sprout

grow 35  
germinate 161  
off-spring 167  
\*vegetable 365, 367\*  
expand 1 94  
-from 154

- item 35, Increase, row 2: V increase, augment, add to, enlarge; dilate &c 194; grow, wax, mount, swell, get ahead, gain strength; advance; run \*shoot\*, shoot up; rise; ascend &c, 305; sprout &c 194.
- " 129 Infant, row 5: scion; sapling, seedling; \*bud\*, tendril, \*shoot\*, olive-branch, nestling, chicken, duckling; larva, caterpillar, chrysalis, etc.
- " 160, Weakness, row 4: weakling; infant, &c, 129.  
(Delete &c 129 and insert as below: )  
weakling; infant; \*mite, tot, little one, slip, seedling, tendril, shoot, whelp, pup, lamb; infantile, puerile, babyish; newfledged, callow.\*
- " 451, Thought, row 5: V think, reflect, reason, cogitate, excogitate, consider, deliberate; bestow thought upon, bestow consideration upon; speculate, contemplate, meditate, ponder, muse, dream, ruminare, \*run over\*; brood over; animadvert, study; bend the mind, apply the mind &c, 457; digest, discuss, hammer at, weigh \*prove\*, perpend; realise, appreciate, \*find\*; fancy, &c 515; trow.  
row 9: occur; suggest itself; come into one's head, get into one's head; strike one, \*strike one as\*; \*be\*; run in one's head, etc.
- " 454 Topic, row 1: N food for thought; mental pabulum; hobby, interest, &c 451\*  
row 2: subject, subject-matter; theme, \*question\* topic, thesis, etc.
- " 455, Curiosity, row 4: Adj: curious, \*interested\*, inquisitive, burning with curiosity, etc.
- " 457, Attention, row 1: attention; mindfulness &c, adj.; intentness; thought &c, 451; advertence; observation; consideration, reflection; heed; particularity; notice, regard &c, \*interest, concern\*; circum-spection, &c, 459; study, scrutiny, etc.

row 2: catch the eye, strike the eye; attract notice;  
catch, awaken, wake, invite, solicit; attract,  
claim, excite, engage, occupy, strike, arrest,  
fix, engross, absorb, rivet the attention, mind,  
thoughts; \*strike one, strike one as\*, be  
present to, uppermost in the mind.

item 461, Inquiry, row 3: sifting, calculation, analysis, \* specification\*,  
dissection, resolution, induction.

" 780, Property, row 8: money, &c 800; what one is worth; estate and  
effects: \*share-holdings, business assets, business

" 894, Courtesy courteous, polite, \*attentive\*, civil, mannerly,  
urbane, etc.

6. What is claimed for the Thesaurus-procedure is the following:

(i) It is a planned procedure for producing idiomatic translation.  
When the translation fails, it is possible to see why.

(ii) Translation-trials made by using it throw unexpected light on the  
principles of construction of a Thesaurus. They should, therefore, yield  
information which will facilitate the construction of a Thesaurus strictly  
compiled on statistical data for scientific M.T.

(iii) On this procedure, the only bi-lingual dictionaries used are word-  
for-word pidgin dictionaries. Nearly all the dictionary-making is done in the  
target language, in which the work of compiling the Thesaurus, however labor-  
ious, need only be done once, since the Thesaurus will transform the mechanical  
pidgin produced from all languages.

(iv) The Thesaurus procedure uses previous M.T. results, which have esta-  
blished the high degree of intelligibility which can be reached by a mechanical  
pidgin, while at the same time keeping open the possibility of further analys-  
ing the input text.

As against this, it will be urged that for M.T. the whole procedure is  
quite impracticable, since no computer could hold a coded Thesaurus. This is  
true, if the Thesaurus were to be actually constructed and kept in being.  
The possibility exists, however, if all the items form lattices, of coding  
merely the chunks of the English language, together with a specification of  
the thesaurus-positions in which each occurs. This presents a formidable  
coding problem; but, with modern techniques of compressed and multiple coding,  
not an impossible one. Once idiomatic M.T. is what is aimed at, a problem  
of comparable order would be presented by the necessity of coding, say, the  
two-volume concise Oxford Dictionary. Current comments on the literature,  
moreover, already make it clear that the commercial world is not going to be  
satisfied with anything short of an attempt to provide multilingual, fully  
idiomatic M.T., since, the better the mechanical pidgin which is provided for  
the commercial readers' inspection, the more impatient the reader becomes with  
the fact that it is not wholly intelligible and correct.

Margaret Masterman.

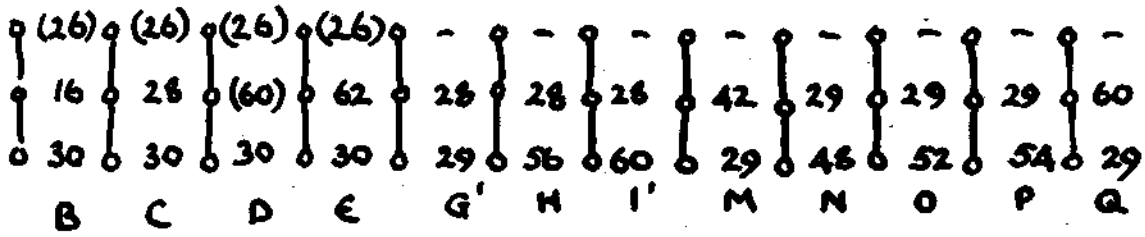
Cambridge Language Research Unit,  
October 6th, 1956.

# A STRUCTURAL DIAGRAMS OF LATTICES

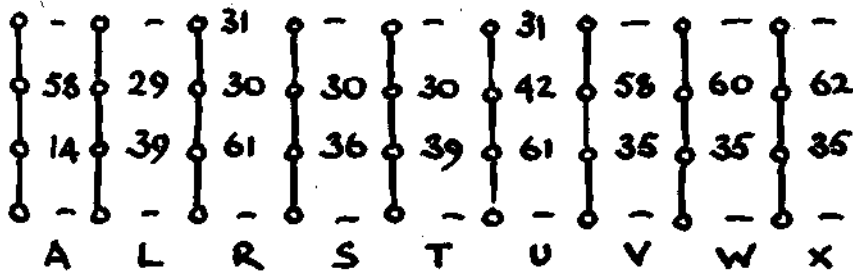
## TWO ELEMENT CHAINS:-



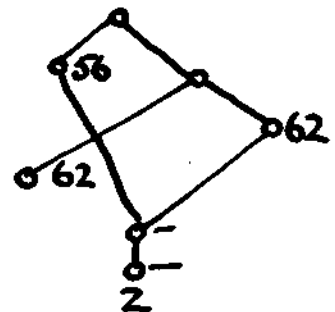
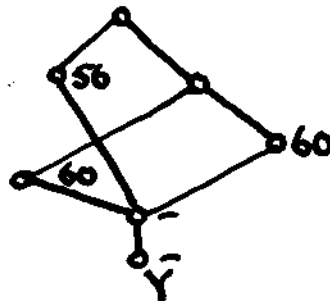
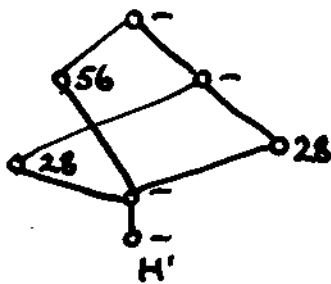
## THREE ELEMENT CHAINS



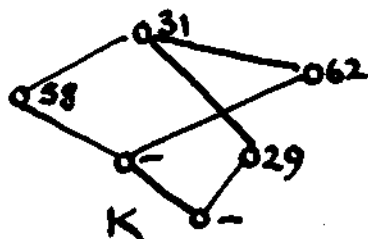
## FOUR-ELEMENT CHAINS



## INVENTOR TYPE LATTICES



## DYADIC TYPE LATTICES:



## COMPOUND TYPE LATTICES

