

R.H. Richens : Translation in the NUDE

Karen Sparck Jones
Computer Laboratory, University of Cambridge

This paper in its final form appeared in *Early Years in Machine Translation*, Ed. W.J. Hutchins, Amsterdam: John Benjamins, 2000, 263-278.

This paper is about the development of R.H. (Dick) Richens' work on MT at Cambridge through the Cambridge Language Research Unit (CLRU). I am not familiar with what he had done before, but he had an early interest in MT and had collaborated with Booth in research written up in 1948 but not published until later (Richens and Booth, 1955). Between then and his involvement with CLRU, he apparently did not do any work on MT, apart from responding to inquiries by journalists and to J.E. Holmström, who was writing a report for Unesco and wanted to say something about the possible automation of translation (Hutchins, 1997). One point of importance is that his formal employment was with the then Commonwealth Agricultural Bureaux: he was Assistant Director, later Director, of the Cambridge Bureau on Plant Breeding and Genetics. The Bureau was responsible for a specialised abstracting service, so Richens would have experienced at first hand the growth of scientific literature which was one of the main stimuli for research on automatic translation and indexing.

Richens did not publish much on MT, presumably because his work on it arose from personal interest rather than through his regular job. His ideas as developed in the Cambridge context are essentially given in Richens (1956a) and Richens (1958).

Richens' interest was in the role of an interlingua in MT. This was a feature of the CLRU's research on MT, figuring in relation to Chinese ideograms, and in the key form of a thesaurus, in the discussions of the informal group including Richens, Margaret Masterman and the linguist M.A.K. Halliday that was the progenitor of the funded CLRU.

1 Richens' key papers

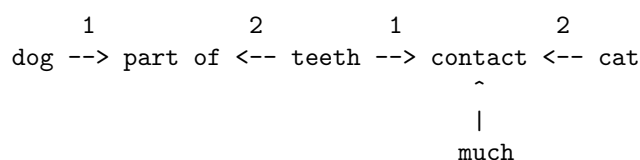
'Preprogramming for machine translation'

Richens' presentation at the Cambridge Colloquium in 1955, 'Preprogramming for mechanical translation' (Richens, 1956a), clearly displays the essentials of his approach in the phrases 'algebraic interlingua' and 'naked ideas'. Thus language meaning would be captured, sufficiently for translation purposes, in expressions constructed from a small set of primitive notions (the naked ideas) using monadic and dyadic operators (the algebra). These are ideas familiar from at least the 17th Century, in the pursuit then of a Universal Character (see e.g. Knowlson, 1975), and Richens' early examples had a rebarbative notation, like some of the 17th Century versions. One of the CLRU's contributions (again following, though doubtless unknowingly, 17th Century precedent) was to make the interlingua more user-friendly, for the necessary practical purpose of making a dictionary.

In the Colloquium paper, preprogramming means designing a general-purpose translation procedure independent of specific machines. Richens was clearly much influenced by the need, in thinking about how translation should be done, not to get prematurely bogged in the particular detail associated with individual punched card operations that mechanisation then typically implied. He also saw designing an MT procedure as accommodating the right sort of distribution of

effort, or ‘companionability’, between man and machine. Thus he considers the levels of processing required for something worthy of the name ‘MT’. They involve not only the lowest level of mechanised dictionary lookup (cf Richens and Booth, 1955) - which he treats in a fairly sophisticated way, though noting that direct translation via a dictionary delivers no more than ‘mechanical pidgin’ - but also syntactic and semantic operations. Richens’ approach to syntax is via pattern matching on word class sequences, using a word-class sequence dictionary: this would consist of pattern-action rules, again for direct translation between two languages, so the actions might e.g. rearrange, delete, or insert material. He clearly envisages a constituent-structure approach with pattern matching applied in cycles, with progressive reduction to a final pattern spanning the whole input.

The focus of the paper is, however, on the ‘more difficult’ semantic processing, for which he proposes what he refers to as a more fundamental approach than sorting out word ambiguity by looking for preferred semantic indicators, say on relative frequency grounds. This is to use ‘an interlingua in which all the structural peculiarities of the base [i.e. source natural] language are removed and we are left with what I shall call a ‘semantic net’ of ‘naked ideas’.’ In such a net ‘the elements represent things, qualities or relations ... [and] a bond points from a thing to its qualities or relations, or from a quality or relation to its further qualification’. At this point, however, Richens does not proceed beyond a few informal examples, e.g. saying that “Dog bites cat” can be represented as



Thus while claiming that a ‘semantic net ... represents what is invariant under translation’, he has nothing concrete to say about how these nets are to be derived from a source text, beyond noting that this is a difficult problem and suggesting a ‘jigsaw analogy’ where ‘each word has a number of semantic properties [like the protruberances on jigsaw puzzle pieces] which fit in with some words and not with others’. In particular, Richens does not discuss the relation between syntax and semantics in detail, and it is not quite clear whether he sees the derivation of a net as following prior syntactic operations or as replacing them. He does however draw attention to the issue of restraining analysis, i.e. not necessarily supplying information missing from the input text: for instance inferring from the occurrence of “solution” in a sentence that something needs to be supplied about a solvent. Thus he also allows that ambiguity may be carried forward from source analysis to target output.

Richens is quite clear about the interlingual potential of his nets: with an interlingua of sufficient expressive and logical power, translation can be much more efficient than by the direct method as he presented it for syntactic processing. As he says in relation to his nets, ‘if the elements (ideas) are replaced by letters with an ideographic significance only, we have in fact an ideographic algebraic script with obvious potentialities for machine translation work’. Finally, he notes that approaching translation as a scaled process with levels of operation suggests that translation be treated as a limiting case of abstracting, so semantic nets could be applied in the latter, for the selective transfer of information, as well as in the full transfer sought for the former, and perhaps with more practical utility given the growth of the scientific literature.

Richens continued to work on these ideas. His next paper was a CLRU working paper. The brief abstract on work in progress given as Richens (1956b) refers to the development of the ‘notational interlingua ... constructed so as to represent the ideas of any base [source language] passage divested of all lexical and syntactical peculiarities; for which reason it is called Nude.’ The interlingual ‘words’ (i.e. expressions) use letter codes for the elements, i.e. for basic ideas like *plurality*, *animality* or *negation*, which are combined via monadic or dyadic relations indicated by different punctuation symbols (which are rather opaquely described). The abstract indicates the relation between syntactic and semantic processing: initial lookup in the base-interlingua

dictionary for ‘semantically significant [subword, word or multiword idiom] “chunks” ’ is followed by syntactic conversion from natural to Nude syntax using the class-sequence method described earlier; and semantic operations are then applied, using the dictionary information, to establish the correct semantic interpretation. What Richens appears to have had in mind here, developing the jigsaw analogy mentioned in Richens (1956a), is the application of some sort of semantic selection restrictions: thus he refers to the ‘interaction entries’ in the base-Nude dictionary. He seems to have envisaged this semantic processing primarily as directed towards word-sense disambiguation, and he apparently did not consider the problems that might arise in the previous stage if the syntactic structure of the input text could not be determined, and when Nude syntax had to be provided *before* Nude semantics. The generation of the target language output is envisaged as involving the same operations, given a Nude-target dictionary.

The full working paper (Richens, 1956c) to which the published abstract refers is interesting primarily in giving substance to Richens’ ideas through a series of worked examples, illustrating his view that ‘only one programme is envisaged for translation between any two languages’, along with a dictionary for each source and target language. The flavour is rather different from that of the abstract, with more emphasis on syntax and less on semantics. Thus the examples show analysis from Japanese to Nude for a brief passage, beginning with character normalisation, chunking and dictionary lookup for morphological and category information. Parsing, i.e. cyclic reduction of Japanese to Nude syntax, is achieved by matching input sequences against word-class sequence entries in the Japanese-Nude dictionary. Richens notes that the order of sequence types (which can imply word reordering) is important: he uses the order qualifier-noun, noun-noun and noun-verb, prepositional phrase, coordinate nouns, adverbial phrase-verb, but without firm commitment. Semantic processing, dealing with dictionary-based ‘interactions’ between text elements is referred to only very briefly, citing Richens and Booth (1955) and Richens (1956a). One such illustrative interaction chooses the sense ‘mature’ rather than ‘tie’ for one chunk given the presence of another chunk meaning ‘fruit/grain’. Generation from Nude to English, German, Latin and Welsh is then shown, following the same step-by-step process and applying the same constituent sequence ordering as in analysis. Within its limits the whole is a tour-de-force, but it shows only the successful steps, and there is no information about the other possibilities considered and rejected. It is worthy of note, however, that while the list of Nude elements given in the paper has only 19 members (e.g. *becoming, straight, near, textile*), that of categories has 32 members.

‘The thirteen steps’, another workpaper apparently written a little later, but undated, (Richens, 1956?), is a further attack on the business of specifying the processing steps and dictionary data required for MT. It goes into considerable detail about e.g. data management, again showing how early MT researchers had to struggle with the nitty gritty of storage allocation, and works through the thirteen translation steps from Latin to English and back again for an example sentence, with much illustration of dictionary codes. The paper discusses syntactic operations a little more fully than in the previous case, but parsing is still confined to immediate rather than remote (i.e. discontinuous) category sequences. Matching, applying the pattern dictionary sequences to the text, is from right to left, reducing when possible, with alternative categories considered only if matching fails. Interaction is illustrated only for syntax (the Latin ‘persecut-us est’, initially treated as passive by combining ‘-us’ and ‘est’, is reinterpreted as non-passive because ‘persecut’ is a deponent verb). Semantic analysis is rather summarily covered by the statement that the form of testing to deal with chunks with alternative interlingual equivalents is uncertain, but ‘it may be possible to compare the ambiguous chunk with all chunks to which it is bonded. The alternative chosen is that with the greatest number of semantic components in common with the chunks bonded to it. It may also be necessary to compare ambiguous chunks with the sum total of semantic elements from the preceding sentences of the same paragraph.’ However after this, Richens simply continues by saying that the first half of the translation process is now complete and goes on to the generation steps.

This workpaper, again somewhat surprisingly in view of Richens’ earlier remarks, has the same focus on morphology and syntax rather than semantics as the previous one. But it is clear that Richens found it was necessary to sort out how the important parts of processing not requiring semantics, or providing the platform for semantic operations, should be done, before elaborating

on semantic operations. In his later papers he concentrates much more on semantics and the interlingua.

‘Interlingual machine translation’

Richens’ 1958 paper, ‘Interlingual machine translation’ (published, it may be noted as a sign of the times, in the first volume of a new computing journal) further develops his ideas and, more importantly, provides rather more concrete information about his interlingua.

Thus as before, while he allows for practical considerations, his primary argument for using an interlingua is theoretical: ‘Linguistic and translation problems are ... more clearly and usefully formulated in terms of a standard language, devised, as Wittgenstein (1922) once suggested, to mirror the logical multiplicity of the state of affairs which is being represented. Thus the twelve English terms “stallion”, “bull”, “ram”, “mare”, “cow”, “ewe”, “colt”, “calf”, “lamb”, “horse”, “ox”, “sheep” can obviously be replaced by three terms for the animal species and terms, respectively, for sex, masculine, youth and contrariety. It is redundant to allocate a term for female, which can be defined in terms of *sex*, *male* and *contrariety*. If preferred, feminists could define “male” in terms of *sex*, *female* and *contrariety*, but it is not possible to dispense with both *male* and *female*’ [notation modified to better distinguish natural from interlingual terms].

Richens’ interlingua consists, as already mentioned, of a network of bonded semantic elements. These are of two types: ‘a limited number of primary elements, of the order of 50 to 100’, representing such fundamental ideas as *exist*, *contrariety*, *cause*, *past in time*, *animal*, *perception*, *desire*. Many concepts can be defined in terms of these primary elements. Thus “giving”, “receiving”, “donor”, “recipient”, “gift” can be defined in terms of ... *cause* and *pertain*.’ But when this is not possible, Richens invokes ‘arbitrarily numbered subcategories’. Thus while “canine” is not sufficiently defined by *animal* and *pertain*, with *dog* defined as subcategory *animal 359* “canine” is completely definable. (There are obvious analogies here with Katz and Fodor (1963).)

Again, the ‘bonds linking the primary semantic elements are of two types, ... homogeneous and heterogeneous. Homogeneous bonding corresponds to the usual idea of qualification. However, no distinction is made between the qualifier and the qualified: we do not distinguish between “black dog” and “canine blackness”. ... The heterogeneous bond is required for dyadic relations, exemplified by most prepositions and transitive verbs.’ Richens emphasises that ‘at the interlingual level, there are no units corresponding to the word ... and no distinctions between elements ... corresponding to parts of speech. The only distinction is that some semantic elements require heterogeneous bonding while others do not.’ Clearly, also, in contrast to natural languages, ‘linear ordering has no significance, all [interlingual] syntactic relations being expressed by the linkages of the bonds’.

The illustrations that Richens gives of his formula are, as mentioned earlier, very reminiscent of some of the 17th Century notations: though he was of course envisaging the notation as an economical, machine-internal one, and as he says, ‘in the construction of mechanical dictionaries, mnemonic catchwords are more useful’. Thus a homogeneous bond between two elements is indicated by attaching the same superscript to each element, so $\mathbf{b} = \mathbf{animal}$, $\mathbf{v} = \mathbf{male}$ and $\mathbf{e} = \mathbf{emotional\ awareness}$, the code for the statement that animals have feelings would be represented as

$$\begin{array}{l} \mathbf{a} \mathbf{a} \\ \mathbf{b} \mathbf{e} \end{array} .$$

Heterogeneous bonds are more complex, since the relation is distinguished by superscript 3, while the terms it connects carry the same superscript followed by 1 and 2: thus the code for ‘somebody thinking about animals’ (a zoologist? or for the sentence ‘Someone is thinking about animals?’), where $\mathbf{m} = \mathbf{mankind}$ and $\mathbf{t} = \mathbf{cognition}$ is

$$\begin{array}{l} \mathbf{a1} \mathbf{a2} \mathbf{a3} \\ \mathbf{m} \mathbf{b} \mathbf{t} \end{array} .$$

Richens next discusses the application of the interlingua in translation. The account here is much as in Richens (1956a), involving an array-based procedure where columns starting from the first with the input words, followed by ones with their dictionary-supplied (word) grammatical, syntactic, semantic and target equivalent information, form the basis for the derivation of further columns: this is the reductive processing described earlier, that by operations on the columns or rows removes redundant information and builds up the input interpretation. Richens notes that the input language syntax operations via sequences of word classes is much like a procedure of Yngve's, though the latter uses it for word reordering. Richens' account of the important semantic stage of input interpretation is slightly fuller than the earlier ones, namely that it uses three types of process, all exploiting the primary interlingual elements. The first type considers 'semantic congruence relations', for example that the element *cognition* can be applied to anything, but only *mankind* can be applied to it. As Richens puts it, 'This is only a formal statement of the fact that anything can be thought of, but only human beings can think.' The second type are 'precise semantic determinations', as in the fact that 'The English "last" is likely to be an appliance only if shoe-making is concerned, and not even then if the interaction "stick", "to", "last" is demonstrable.' It is not quite clear how Richens would achieve such refined interpretations, though it may be that they would involve the invocation of his subcategories. The third type are what he calls 'diffuse semantic determinations', involving the comparisons moving gradually outwards from immediately bonded elements, a form of processing he mentions as studied in the CLRU's thesaurus research. It has to be said that Richens' full description of processing gives no more material detail than this summary: thus the points made earlier suggest that he did not really recognise (even if he allowed for a quality and precision loss in translation) that an essentially word-driven approach, much focused on lexical ambiguity resolution, could get really bogged in syntactic structure determination.

'Tigris and Euphrates'

Richens' last major paper on language research, 'Tigris and Euphrates - a comparison between human and machine translation' (Richens, 1959), is an extensive, somewhat reflective paper, given at an important international meeting, the 1958 NPL Symposium on the Mechanisation of Thought Processes. It is interesting in capturing the flavour of the period, for example in relation to the purely practical problems encountered in automation, as well as in presenting Richens' own views within a larger framework, though it does not add to the detail on his approach.

In the first part Richens focuses on the symbol categories involved in language use, and on translation as a symbol transformation process via (inter)mediate categories, notably those of lexicon-grammar, syntax, and naked ideas. Translation can work if the categories of a given type are sufficiently similar for two languages; or, in the special N category case for naked ideas, have common unique indicata, to the degree of information resolution that is functionally adequate: Richens is not an absolutist. Richens illustrates, with many apposite examples from different languages, the way that language symbol strings are characterised, by a decomposition to category labels with bonds tying (and implicitly ordering) related elements. Translation is thus a formal operation mapping between natural and symbolic language expressions or between expressions in two symbolic languages. He works through the whole sequence of transformations for the sentence "She came to", emerging with an N form that can be crudely paraphrased as

[(Xref female one)(become conscious)(before now)]

Richens does not claim humans explicitly or fully deploy the symbolic processing he invokes for MT; but he does claim these symbol types do show up sufficiently for it to be clear they are real. Further, with respect to the N category, he maintains that it is exceptional for 'human translation to proceed without any recourse to symbols of this degree of generality'.

In the second part of the paper, comparing human and machine translation, Richens considers not only the difficulties of automating individual operations that humans find easy, like word segmentation (cf also Richens and Halliday, 1957), but those that arise because practical machine limitations mean that symbol type mappings have to be done in separate stages. As a consequence,

given that ambiguity resolution is the major requirement at each step, while humans can range at once over different types of information, machines are confined to one at a time. Thus even while as much resolution should be done at each stage as possible, much will inevitably be left for later determination. As Richens puts it in the case of syntax, machine operations are stereotyped, being confined to the use of relations between grammatic categories. This is exemplified by his own use of a syntax pattern dictionary. However while syntactic analysis resolves many ambiguities, there is a ‘residuum’ for semantic analysis. This is the real challenge, where ‘the powers of the human translator are seen at their subtlest’ and where for MT, Richens’ faith in the interlingual approach is evident. Thus he asserts that ‘no method of solving the problems connected with multiple [word] use without recourse to N-symbol analogues has been devised’. However he accepts that several different types of semantic procedure will be needed, determining ‘semantic congruence’ between bonded N symbols, identifying special collocations of lexico-grammatical symbols, or invoking shared ‘semantic fields’ as in the CLRU’s thesaurus work.

Richens conclusion is that while MT has no immediate prospect of rivalling the all-round performance of the human translator, it may be more efficient for some processes, and occasional produce superior translations.

In this historical context, it should be mentioned that Bar-Hillel, as commentator, while noting that Richens and Booth (1955) was ‘the first serious contribution to machine translation altogether’, attacks Richens’ foundations, not just from a philosophical point of view but as a professor dissecting an inadequate student essay. Richens’ defence is that he is trying to present practically-oriented work, and in using terms like ‘indicatum’ to this end, not to make deep philosophical claims. Bar-Hillel in particular maintains that Richens has not produced any real argument to show that naked ideas will meet with better success in the 20th century than they did in the 17th: only a complete system can do that. But Richens knew that the proof of the pudding would be in the eating.

Though Richens’ views about the form and application of an interlingua were one ingredient the CLRU’s initial ferment of ideas, the (NSF-) funded work at CLRU was primarily devoted to research on the use of a thesaurus for MT, taking Roget’s *Thesaurus* in its English Penguin Books edition as an exemplar and practical experimental tool. This work included studies of automatic word sense selection (with primitive punched card machine operations as successors to manual simulation), formal modelling using lattice algebra, investigations of thesaurus use for a range of NLP tasks, not just translation, and experiments in automatic thesaurus construction. The way that Richens’ ideas about the form and role of an interlingua were developed at the CLRU was largely a response to the perceived limitations of the Roget-type thesaurus as a vehicle for language processing, so it is helpful, before further elaborating on the CLRU’s use of NUDE, to summarise the CLRU’s experience with a thesaurus (cf Sparck Jones, 1992).

2 The thesaurus

The CLRU had advocated, from the mid fifties, the use of a thesaurus as a sense selection device for language processing. The concepts represented by word classes in a thesaurus could be taken as a set of primitives for categorising word meanings. The classificatory *heads* in Roget’s *Thesaurus*, for instance, supplied a set of 1000 such primitives which could make quite refined distinctions between the senses of any given word, as the index in the printed thesaurus clearly shows; and at the same time, the classes grouped words with similar meanings. However, while it was convenient to take the class labels as the names for semantic primitives, the CLRU always saw class concepts as ostensibly defined by the sets of class members, and classes themselves as based on, and hence constructible from, the observed behaviour of words in text. As noted earlier, the CLRU argued that such semantic classifications had a role in different language processing tasks, for example indexing as well as translation; and on the reasonable basis that humans deal with the same or similar worlds whatever their specific language, the CLRU also argued that the primitive concepts of the thesaurus could offer a useful interlingua.

As with Richens’ version, there was nothing especially original about this form of interlingua,

whether seen from the perspective of a Universal Character or from that of the search in linguistics for language universals. The novelty in the CLRU's work was trying to put the thesaurus to work, in real computational practice: for MT this implied using it both in identifying source language word senses and in choosing target language equivalents (Masterman et al., 1957/1986). To that extent whether such categorial primitives are out there in the world, or in here in the head, is immaterial: the issue was whether some set of primitives could be found that was functionally sufficiently effective for MT. Thus the analogy is with the practical lexicographer who, though he lists three senses for a word, does not maintain the number three is absolute, only appropriate for the purpose to hand. One consequence was that thesaurus characterisations of word senses were never seen as exhaustively capturing meaning, only as capturing enough of it for generally appropriate sense and equivalent selection.

The CLRU, in trying to apply Roget as a tool, nevertheless encountered the same sorts of problem at a practical level that were, a little later, brought as theoretical charges in the linguistics literature against Katz and Fodor's semantic markers (Katz and Fodor, 1963; Bolinger, 1965). One of the most bothersome problems was that representing a word sense with a single class concept of the thesaurus kind, even with the large set of Roget categories, failed to capture - or certainly failed to express - the fact that word meanings can have structure. This applies whether the word meanings are internally more complex than simple category allows, e.g. 'basket' is a more complex notion because of the way baskets are made than the tag RECEPTACLE indicates; or because complex external relationships are implicit in word meanings, e.g. (to) 'transport' implies at least a transporting agent. Even with finer categories, it is necessary to unpack relational structure in order to be able to establish the cross-text connections that will resolve ambiguity.

Richens' style of interlingua was precisely intended to deal with this problem, and it was therefore given a working implementation for CLRU research. Thus it was applied in an extensive exercise in dictionary making, nominally for Italian (for reasons now unclear but which probably stemmed from a scheme for translating sample texts, including an Italian botanical one, from various languages) but actually for the English equivalents of Italian word (senses). The description which follows is for NUDE as established for, and through, this exercise. Thus the basic set of primitives, and the syntax for interlingual expressions (i.e. word definitions), are due to Richens, but minor developments of the primitive set and conventions about formulae are the result of the way the interlingua evolved and was then consolidated and finally frozen for the sake of consistency across dictionary entries, during the lexicon project.

3 The interlingua NUDE

The material below is a simple account taken from an internal CLRU memo, 'A note on 'NUDE'' written by the present author in 1963. Analysis and observations on the CLRU's experience with NUDE are in the next section.

NUDE had 51 elements, i.e. semantic classifiers, namely

ASK BANG BE BEAST CAN CAUSE CHANGE COUNT DO DONE FEEL FOLK FOR GRAIN
 HAVE HEAT HOW IN KIND LAUGH LIFE LINE MAN MORE MUCH ONE PAIR PART
 PLANT PLEASE POINT PRAY SAME SELF SENSE SIGN STUFF THING THINK TRUE
 UP USE WANT WHEN WHERE WHOLE WILL WORLD

plus the special element NOT, which could be applied only as a direct operator on an individual element. The elements had Anglo-Saxon monosyllabic names for independent reasons.

There were two connectives, ':' and '/'. The ':' connective defined a symmetrical, mutually modifying relation between two elements, so in 'A : B' A and B qualify one another, The '/' connective defined an asymmetrical relationship between two elements A and B, so in 'A / B' A operates on B. In addition, to ensure that formulae were not ambiguous, brackets were used; more specifically a binary bracketing principle was applied, so formulae could combine pairs of elements, a pair consisting of an element and a bracket group, or a pair of bracket groups. There was thus

a definition of a well-formed NUDE formula, which consisted either of an element, or of a pair of well-formed formulae linked by a connective and surrounded by brackets, as in

A / (B : C)
 (A / B) / C
 A : (B / (C : D))

The practical constraints of the punched cards used for the Italian dictionary placed an upper limit of eight elements, and of six for one half, on formulae. This also ensured some consistency in the depth of meaning analysis.

The lexicon experiment showed that there were in fact some important (i.e. recurring) notions that de facto functioned as primitives, for which there was no existing NUDE element. This led to the provision of so-called prototypic formula, used in a fixed manner as if they were single elements. They included

IN : THING (container)
 DO : FOLK (custom, ritual)
 DONE : CHANGE (event)
 CAUSE / HAVE (give)
 CHANGE / WHERE (move)
 HAVE : (SIGN : STUFF) (money)

and others for *communicate, tool, -ology, condition or state of affairs, city or country, group of people, collection of things, necessity or must, one another, symbol or picture.*

With this apparatus, dictionary entries like these were constructed for the Italian dictionary:

‘now’ (POINT : SELF) : WHEN
 ‘to distress’ (CAUSE / (SENSE : NOT PLEASE)
 ‘entirely’ WHOLE : HOW
 ‘to affirm’ CAUSE / (HAVE ((TRUE : BE) : SIGN))
 ‘to hurry’ CAUSE / ((MUCH : CHANGE) : WHERE))
 ‘comfortably off’ (MUCH : (HAVE : (SIGN : STUFF))) : HAVE
 ‘agent’ ((NOT FOR / SELF) : DO) : MAN
 ‘to join’ (CAUSE / (BE / PART)
 ‘sharp-witted’ BANG : THINK

4 Commentary

The Note emphasises the properties of the interlingua as Richens originally viewed it. A NUDE formula was regarded as a definition of word meaning, though it was not (given the small set of primitives) normally a unique one. Thus a NUDE dictionary would give a two-stage specification of a word, with the formula at the upper level, and a list number indicating the precise word in a list of synonyms with the same formula on the lower. Naturally, the elements as semantic classifiers would function just like thesaurus ones, so e.g. the recurrence of common classifiers over a text could be used as a device for resolving lexical ambiguity. (The same principle could indeed extend to subformulae.)

The further, language-like, property of NUDE is very evident when whole formulae are considered, where it is easy to see elements taking on the behaviour of nouns, verbs and adjective: thus the connective ‘:’ was seen as joining elements in the way an adjective and noun could combine, and ‘/’ as linking a subject and verb or verb and object. This language property was further emphasised when the NUDE formulae for the words in a sentence were laid out. NUDE had a cheerful pidgin character, precisely as Richens’ view of it as a basic language implied. Thus this author and Roger Needham are still wont to refer to any VIP (or ‘big wheel’) as a (MUCH : UP) : MAN.

Some of the problems with NUDE also emerged in the lexicon effort, as appears in the Note. For example, some elements developed a conventional syntactic function, so **HOW** came to be a standard marker at the end of a formula for an adjective. This made it difficult to apply as a notion in its own right for the concept of manner. The same applied to **DO**, which if used conventionally for verbs could not also be used to indicate activity as an abstract notion. There were also problems in providing any adequate definitions for some types of words, notably prepositions, where the NUDE elements could not make even major sense distinctions. Again, just as with Ogden and Richards' Basic English, or the restricted definitional vocabulary used in the *Longmans Dictionary of Contemporary English*, the generality of the elements meant they themselves developed a wide range of meanings and it was a major effort, in the dictionary making, to maintain consistent practice across entries. As an example, though the ':' connective was symmetrical, it acquired a 'modifier-head' ordering, with the dominant element in the formula in rightmost position (as in the VIP example). (It should be noted that Richens' original idea of subcategory lists was accepted, though it was not very rigorously implemented, being regarded as a subordinate routine matter.)

The syntactic structure that NUDE imposed on definitions was a clear advance, for finer NLP tasks, on the simple set-based semantic definitions provided by the thesaurus. The Italian Dictionary was built up to over a thousand entries (probably more), with all members of the CLRU undertaking a daily quota and manifestly showing they could all speak NUDE, albeit with variations in idiolect. The work with the primary version of NUDE just described also prompted research on a number of alternatives and extensions, for example organising the elements into a hierarchy, incorporating element order into formula interpretation, and so forth, as illustrated in Masterman (1962), where an attempt was made to combine Roget's Thesaurus and NUDE.

5 Later developments

The research on NUDE and interlinguas was actively pursued at the CLRU in the latter part of the fifties and first years of the sixties. Richens was an involved participant, especially in the earlier period, though he was not a CLRU employee. His personal style, an engaging combination of personal modesty, practical orientation, and intellectual conviction, made him a persuasive advocate, and his claims for his interlingual approach were justified by the fact that the CLRU was able to carry through a body of research with it. Overall, however, the work was less directly productive within the CLRU than might have been hoped or expected. Thus it was difficult to proceed from building the dictionary to actually engaging in translation using it (even supposing the NUDE semantic definitions were supplemented with other information, e.g. about word classes, as Richens envisaged). Some of the reasons were extrinsic. Richens himself became increasingly detached, largely through the pressure of his CAB work. Funding requirements pulled the CLRU work into other directions, notably towards information retrieval, where simpler definitional schemes were more appropriate. Other, more seductive topics and paths appeared, e.g. work on text prosody, which, when Margaret Masterman as Director pursued them, naturally had an influence on the CLRU's research focus.

However the main weakness of the CLRU's attempt to use NUDE as a vehicle for translation (as, similarly with its attempt to use the thesaurus), was a failure to get the whole processing package together, and specifically to address the interaction between syntax and semantics. Though the CLRU's initial concentration on semantics rather than syntax as the real nut to crack for unrestricted MT was commendable, there was never, especially after Martin Kay's departure for the US, enough effort devoted to the treatment of syntax and the combination of syntactic and semantic processing. The CLRU's work on MT was a car without a powerful enough engine. In principle, NUDE reached far more towards syntax than the simple thesaurus; and as mentioned earlier, Richens proposed mechanism for translation took account of syntactic information. But as the account of his array-based procedure makes plain, it was still viewed essentially as a word-level operation. There was little recognition in Richens' accounts of what the need to treat whole (larger) constituents as units would actually involve. Thus though NUDE itself could express relational structures, and could do this in principle in a way that could extend from individual words

to multiword structures, the CLRU never really tackled the details of how this should be done using syntactic information, and of how the case-type relations that NUDE involved or implied could be built from more specific, conventional morpho-syntactic information.

This is partly attributable to the then current style of syntactic description, namely transformational grammar: this was exceptionally unsympathetic to lexicon-driven processing and to semantic relational structure. Moreover, while some practical MT work in the sixties was undoubtedly more lexicon-driven, it was squelched along with MT research in general in 1966.

There was, however, another explanation for the lack of progress in what would have seemed the obvious direction, by attending to the use of conventional syntax. This was Masterman's drive to parse text semantically, using message patterns. Thus given the need to shape and focus the interaction between lexical semantic formulae, she began (in Masterman, 1962) to develop the idea of 'semantic shells', or basic message forms, which would provide the anchors, in text words, for text-spanning operations with the formulae. By Masterman (1966) these shells had become three-term 'templates', which she associated with the natural prosodic phrasings of utterances. These phrasings would, moreover, fit meta-patterns of 'semantic squares', determining basic connection patterns between pairs of templates (much as in the local focus-determining rules of later NLP work). Throughout this work Masterman specifically relied on NUDE. Thus she defined templates as NUDE-type formulae, so when templates were mapped onto (i.e. instantiated for) a text, the entire NUDE entries for the words involved would be invoked to support the required semantic disambiguation and structure capture. Masterman adopted, however, such an aggressively fundamentalist approach to this whole pattern determination operation, and so resolutely eschewed help from syntax, that she was never able to carry her ideas into effective computational practice.

Thus while at the CLRU itself the necessary development of NUDE as a mechanism for translation was not carried through, this was done by Yorick Wilks in work begun at the CLRU but finished at Stanford (Wilks, 1972). Though Wilks did not approach syntax in a conventional way, he made use of syntactic information in text analysis. More importantly, in order to identify phrase and clause level structural patterns, via templates and 'paraplates' that were defined by NUDE elements and into which individual NUDE-type word formulae could be slotted, Wilks had also to develop a much more complex view of element categories, word formulae, etc, as well as to supply a battery of patterns. He had to make the notion of formula head quite explicit, and define and express case relations (thus his work had much in common with Schank's later Conceptual Dependency approach). With the degree of elaboration that he gave it, Wilks was able to show, in computational experiments, that the type of interlingua that Richens had adumbrated could form the basis, albeit not the sole basis, for NLP. His research was followed up in Boguraev's work (Boguraev, 1979), in which a more wholehearted use of conventional syntax was taken as an underpinning for the derivation of analyses, and which in turn delivered more structured case-based sentence representations than Wilks' own system. The persistent, and valuable, line in NLP that exploits semantic patterns, and which is now a routine strategy for restricted task and domain applications, thus owes a debt to Richens.

It is not clear what Richens himself made of Masterman's ideas as she developed them in the sixties, or how far he knew of the details of Wilks's computational experiments. His habitual expression was always a little quizzical. He also had other interests, and made another, rather different contribution to the common heritage. As a botanist, he was especially interested in the elm tree, and published a classic, *Elm*, on it in 1983. This now has especial poignancy since the ravages of Dutch elm disease make the illustrations of individual fine trees reminders of another, lost age. Richens himself died in 1984.

6 References

- Boguraev, B.K. *Automatic resolution of linguistic ambiguities*, PhD Thesis, Technical Report 11, Computer Laboratory, University of Cambridge, 1979.
- Bolinger, D.W. 'The atomisation of meaning', *Language*, 41, 1965, 555-573.
- Hutchins, W.J. 'From first conception to first demonstration: the nascent years of machine translation, 1957-1954. A chronology', *Machine Translation*, 12, 1997, 195-252.
- Katz, J.J. and Fodor, J.A. 'The structure of a semantic theory', *Language*. 39, 1973, 170-210.
- Knowlson, J. *Universal language schemes in England and France, 1600-1800*, Toronto: University of Toronto Press, 1975.
- Masterman, M. et al. 'Agricola incurvo terram dimovit aratro', Report ML 92, Cambridge Language Research Unit, Cambridge, 1957; reprinted with a new introduction by K. Sparck Jones, Computer Laboratory, University of Cambridge, 1986.
- Masterman, M. 'Semantic message detection for machine translation using and interlingua', in *Proceedings of the 1961 International Conference on Machine Translation and Applied Language Analysis*, London: Her Majesty's Stationary Office, 1962, Vol II, 437-474.
- Masterman, M. 'Semantic algorithms' in *Proceedings of the Conference on Computer-related Semantic Analysis* (Las Vegas, 1965, sponsored by NSF, ONR and USAF), Wayne State University, 1966, IV 1-97.
- Richens, R.H. 'Preprogramming for machine translation', *Mechanical Translation*, 3 (1), 1956, 20-27. (1956a)
- Richens, R.H. 'A general program for mechanical translation between two languages via an algebraic interlingua', *Mechanical Translation*, 3 (2), 1956, 37. (abstract only) (1956b)
- Richens, R.H. 'A general programme for mechanical translation between any two languages via an algebraic interlingua', Report ML 5, Cambridge Language Research Unit, Cambridge, 1956. (1956c)
- Richens, R.H. 'The thirteen steps', Report ML 5, Cambridge Language Research Unit, Cambridge, 1956? (1956?)
- Ritchens, R.H. 'Interlingual machine translation', *The Computer Journal*, 1 (3), 1958, 144-147.
- Richens, R.H. 'Tigris and Euphrates - a comparison between human and machine translation', *Mechanisation of Thought Processes, Proceedings of a Symposium, 1958*, London, HMSO, 1959, 281-302; discussion 303-307.
- Richens, R.H. *Elm*, Cambridge: Cambridge University Press, 1983.
- Richens, R.H. and Booth, A.D. 'Some methods of mechanised translation', in Locke, W.N. and Booth, A.D. (Eds.) *Machine translation of languages*, 1955, 24-46. (apparently originally 1948)
- Richens, R.H. and Halliday, M.A.K. 'Word decomposition for machine translation', *Report on the Eighth Round Table Meeting on Linguistics and Language Studies*, Georgetown, Washington DC, 1957. 79-.
- Sparck Jones, K. 'Thesaurus', in S.C. Shapiro (Ed.) *Encyclopedia of Artificial Intelligence*, 2nd Edition, 1992, 1605-1613.
- Wilks, Y.A. *Grammar, meaning and the machine analysis of language*, London: Routledge and Kegan Paul, 1972.