

Semantic primitives: the tip of the iceberg

Karen Spärck Jones

Computer Laboratory, University of Cambridge *

This paper in its final form will appear in *Words and intelligence: Part II: Essays in honour of Yorick Wilks*, Ed. K. Ahmad, C. Brewster and M. Stevenson, Berlin: Springer, 2007.

Abstract

Semantic primitives have been central to Yorick’s approach to language processing. In this paper I review the development of his ideas on the nature and role of primitives, considering them both from the narrower system point of view and in the larger context to which Yorick himself always referred.

1 Introduction: semantic primitives

(((*ANI SUBJ)((FLOW STUFF)OBJE)((SELF IN)((WRAP THING)FROM)((MAN PART)TO)(MOVE CAUSE))))

I want to revisit Yorick’s 1983 question: “Does anyone really still believe this kind of thing?” (Wilks 1983a). Yorick in 1983 argued for the kind of thing illustrated by the formula above as a semantic tool for resolving what others would describe as awkward syntactic problems. Here, to see what a contemporary answer might be, I will look again at how Yorick’s ideas about “this kind of thing” developed and what they may say to us now.

My focus is thus on *semantic primitives* and their intimate relationship with word sense disambiguation, as well as ambiguity resolution more generally. But semantic primitives are the tip of a large iceberg with natural language processing and its tasks in the upper layers, and the philosophy of language and its ramifications in the lower ones, a continuity illustrated by the connection between primitives, interlinguas for translation, and the language of thought. The iceberg has many other component lumps: discourse structure, and metaphor, for example, and Yorick has always sought to relate his work on automatic language processing with linguistic theory on the one hand and philosophy on the other. Thus Wittgenstein and Quine are invoked as philosophical supports for his position, and language processing, its

*I am grateful to Ted Briscoe for comments.

needs, and his strategies for it are brandished as tools to attack theoretical linguists of the Chomskyan or Generative Semanticist schools. I will concentrate on Yorick as a language processor, so to speak, but also comment more briefly on these other facets of his work.

I will begin with Yorick's first substantial paper, "Computable semantic derivations" of 1965. This presented many of his basic ideas, which were developed, especially for computational experiments, in a phase of research cumulatively represented in Wilks (1972a), with philosophical and theoretical amplifications also illustrated by Wilks (1971) and Wilks (1975a). Then, though much remained the same, there were some changes, especially under the growing influence of research on artificial intelligence, rather than the earlier machine translation: this phase can be roughly taken up to Wilks (1977a), which offers some theoretical discussion to amplify system-centred papers like Wilks (1975b). Some further shifts are signalled in Wilks (1978), and amplifications in Fass and Wilks (1983) and Wilks (1983b), followed by a much more substantial change of direction as signalled by Wilks et al. (1987), and developed in Wilks et al (1996). My focus will be on Yorick's work up to 1983, and I will consider his later machine dictionary work, as represented by Wilks et al. (1987, 1989) and Wilks et al. (1996), primarily as a comment on the earlier ideas.

2 Computable semantic derivations

It is important to put this work of Yorick's in its historical context. Research in automated natural language processing at the Cambridge Language Research Unit (CLRU) (see especially Masterman 2005) began in the 1950s and focused from its very beginning on semantic issues, specifically as presented by machine translation though it also sought to relate translation to the use of language in general. Thus for translation, how were the senses of words in the source language to be identified and characterised so as to guide the choice of words (senses) for the target language? The initial model used a semantic classification - a thesaurus like Roget's - to categorise word senses, and the notion of class recurrence or recurrent class association over a text as a device for selecting individual word senses. Syntax was taken for granted, rather than dismissed out of hand. The thesaurus could be seen as an interlingua, so the selected class characterisations of the words of the source text constituted the device transmitting source word meanings, or at least the essential elements of source word meanings, to the target language generator. Each thesaurus class could be taken to represent, or embody, a semantic primitive.

Such a simple model, however engaging its simplicity, was also manifestly inadequate as a strict interlingual model, i.e. one in which there is no presumption that source word senses, once identified, point straight to specific target language equivalents. It was succeeded, for experimental purposes, by an interlingual language, NUDE (Spärck Jones 2000), which made use of a much smaller set of primitives and characterised word meanings not by individual primitives or, in principle, sets of primitives, but as syntactically structured formulae. The members of the CLRU were sufficiently serious to engage in a non-trivial lexicographic exercise, providing dictionary entries for a general vocabulary using NUDE. But the all-important issues of how exactly the dictionary entries were to be applied in conjunction with syntactic information and, as would certainly be required, to intersentence relationships, had been only partly addressed before Yorick tackled them, and the latter much more, by Masterman, than the former. Yorick's key contribution was to address these crucial, and interrelated issues much more thoroughly and to show, through computational experiments, that he could get

something to work (see, e.g. Chapter 4 in Wilks 1972a). This work led, in turn, to further development of the structure of the ground-level lexical formulae.

Yorick's strategy was to enhance the scope and role of semantic patterns that defined message or discourse structures. The basic patterns, *templates*, exploited fundamental 'actor/action/object' relationships with actor, action and object each defined by semantic primitives that figured as the dominant or head primitive in corresponding lexical formulae. This may seem an obvious idea now, but within the then context of machine translation research, with its primary emphasis on syntactic analysis, was extremely revolutionary. The initial application of these ideas in Wilks (1965) was nevertheless very simple: text was fragmented e.g. at prepositions, and fragments were annotated by template primitive triples that could be mapped on to them. Sequencing and compression rules then helped to select templates and template chains characterising more extended text. These rules embodied obvious ideas about repetitive cohesion, and exploited the full formulae for lexical items dependent on the triples. The general model was of gradual convergence on the sense of a text over successive fragments. The paper describes, in extremely opaque detail, actual codings using such familiar primitives as MAN, THING, WHERE, WHEN, BE, MOVE along with other perhaps less familiar ones like TRUE and SIGN, templates like MAN WHERE BE and MAN KIND BE, and individual word formulae like ((NOT TRUE : LIFE) : (MAN : KIND)) for "ill" and ((LIFE : STUFF) / (MAN / IN) : DO)) for "eat", along with some computational experiments which show, realistically, how much apparatus is required to deal even with tiny texts.

Some general ideas characteristic of Yorick's subsequent work already appear in this early paper. First, working with minimal syntax: template slots were associated with conventional syntactic categories, and the fragmentation made implicit use of natural constituents and their internal word order, but this was far from full-blown syntax. Second, templates characterised the form of expressions of facts, not facts themselves. Third, templates were devices for resolving word, and higher unit, ambiguity in context. Fourth, meaning representations are what remains as the semantic structure of the resolved text as a whole. Fifth, there is no notion of correctness with respect to a text, just as there is no notion of 'the' legitimate senses of a word. Rather, there is what suffices for the task to hand, e.g. machine translation. Finally, and most importantly, that any text processing has to deal, as a matter of course and with its core mechanisms, with new uses of words: there has to be a way of responding to new, extended word uses, and to novel variations on familiar message forms. This is not however illustrated, but stated as an immediate research goal. The unifying notion linking these various features of Yorick's approach is that of *preference*: there are no absolutes about the semantic relations between text components, only preferences, which may be more or less fully satisfied.

3 Preference Semantics: design, development and defence

I have elaborated on Yorick's initial paper, because his work for the next two decades is a response to the exigencies of making it work and, also, an expansion on its theoretical positions, for example in opposition to the Chomskian primacy of syntax. Wilks (1967), for instance, describing experiments in processing question-answer pairs as mini-texts, again emphasises, on the one hand, the notion that his apparatus is essentially heuristic but also, on the other, the two theoretical points, first that one should talk of achieving a text interpretation, rather than identifying the text's correct meaning and second, and significantly, that the basic device

for doing this, the semantic primitives, are themselves only words and behave like words: thus they can be individually ambiguous, though the complex formulae built from them are not.

Wilks (1972a) describes further computational experiments using the apparatus just summarised to process paragraph-length texts (including ones from major philosophical works), and also to deal with new or extended senses of words, essentially by projecting sense characterisations from ‘old’ word senses in the text to the new ones so that template preferences can be satisfied.

Formula and templates, and paraplates and inference rules

However the main development of Yorick’s initial apparatus was to allow for much richer discourse structure, both by dealing with the attachment of subordinate text units, like adjectives, to their nominals and also, much more importantly, by introducing higher-level patterns designed to combine templates. These *paraplates* characterised template pairs connected by case-type relations as signalled by, e.g., prepositions. The paraplates offered a much tighter and more explicit account of the semantic relations between words and word groups. The basic templates of course covered the primary agent/object relations but the paraplates extended case ties to, for instance, TO, LOCA, POSS and GOAL. Deploying all these interpretative components depended on the word formulae which had a complex internal structure with sub-formulae signalling semantic properties of sought case ‘relatives’, on the use of classes of semantic primitives with shared behaviours, and so forth. Thus we now have a formula like ((**ANI* SUBJ) (SIGN OBJE) (((*MAN* SUBJ) SENSE) CAUSE)))) for “singing”. Yorick’s use of case ideas, quite familiar now, and figuring for linguists in Fillmore (1968), were innovative in the computational context.

This extended range of semantic resources was nevertheless still insufficient to resolve ambiguity, and was therefore further supplemented by common sense inference rules, essentially rule-based ways of unpacking formulae to make implicit pattern information explicit in template form and hence available for further matching. Common sense inference does refer to (familiar and recurrent features of) the world but does not operate directly on a specific world model and is primarily expressed in conventional linguistic forms. The new linguistic objects that the inference rules produce help to lead to the interpretation of a text as that given by the densest network of preferences linking one text component with another. This interpretation is manifest in the sequence of selected (and presumably implicitly or explicitly filled out) word sense formulae in the primitive-based meaning representation language.

The details of Yorick’s approach to text interpretation, along with the fact of its computational implementation and some successful or at least plausible processing outputs, were used to buttress Yorick’s theoretical arguments against linguists and philosophers of language in, for example, Wilks (1972b) and Wilks and Schank (1974) as well as Wilks (1971) and (1975a). His theoretical position was equally an up-market endorsement, indeed ground for, his computational practice.

Thus in Wilks (1971) he makes an attack on grammaticality as not a valid independent property of texts: all that counts is meaningfulness, which in turn is having one of several possible interpretations, implying sense resolution in context and hence meaning for a (coherent) text as a whole. Such an interpretation is a linguistic interpretation, in the sense of being in the same or another language, so text objects are alternatives for one another. Thus one might substitute a dictionary definition as a paraphrase for a word in a text, or a translation

of a text. This is the name of the primary or routine meaning game in the Wittgensteinian and Quinean tradition, anti denotational semantics, with even referential ‘bottoming out’ never certain. Similarly Wilks and Schank (1974) argues that linguistic (and we may say specifically semantic) theory is not to do with grammaticality (or decidability) as properties of sentences, and about competence in being correct on this. Because meaningfulness springs from a dynamic process of text interpretation, and in particular because meanings for new word senses have to be constructed, a semantic theory has to be grounded on a substantive notion of performance.

In Wilks (1975b,c) Yorick gives further detailed, and computationally oriented, accounts of his system. The detail, especially on how the various components individually and collectively work, are nothing like as crude as my summary might suggest. The lexicon, with 600 entries, was far from negligible for the time (or indeed some time after). There is a more detailed treatment of translation, into French (see also Herskovits 1973), guided by a surprisingly specific generation apparatus using stereotypes for French words attached to paraplates. However there does not appear, in general, to be any guidance in output from the specific input words, i.e. the formulae etc. are a true interlingua rather than only a vehicle for selecting from alternative output choices for particular input lexical items. One of the puzzles of these accounts, and the relatively limited and repeated examples given in the papers of the period, is how the interpretive system can actually capture some of the finer grain of the input, for example the specific determiners used. One must suppose, since everything is lexically driven, but primitives like THIS may not be sufficient for unequivocal surface determiner selection, that there has to be a perfectly good conventional bilingual dictionary showing equivalent word forms which, together with the final primitive representation, does lead to appropriate translations including, for French, the choice of “le” or “la” as article.

In Wilks (1975c), Yorick also indicates how his basic apparatus of formulae, templates and paraplates, by characterising stretches of text and showing semantic relations, can help to resolve anaphors, for example by unpacking formulae to exhibit case ties. The suggestion that Yorick’s basic system can be extended to deal with such demanding phenomena is all the more surprising given what appears to have been a concurrent non-trivial change to the system underpinnings, namely the disappearance of the conventional syntactic category information that figured in the initial version. Yorick’s claim is that syntax can be done on the fly and (apart from the rather limited syntactic notions marked by SUBJ and OBJE) is, as it were, finessed by the direct semantic relationships given by the templates and so forth. This is all also, as Yorick emphasises in Wilks (1975c), within a philosophical framework that not only, in terms of a controversy of the time, makes him a proceduralist rather than a declarativist about meaning, but more specifically bases processing on an ‘inertial’ or ‘laziness’ principle.

In these accounts of Yorick’s approach, semantic primitives embody generic concepts shared by many words and as such facilitate key text processing operations. Here, and previously, he maintains that his vocabulary of primitives is stable and, while not claiming that his sets of templates, paraplates or common sense inference rules are equally stable and limited, he does claim, while recognising that his actual testing has been rather modest, that the sets that are required to serve language interpretation and generation tasks will be small.

Then again, in Wilks (1975d and 1977a), he explores the wider theoretical context of his work, with particular reference to primitives. Thus in (1975d) he returns to the fact that primitives are only words, albeit in a ‘small’ natural language, and that there is therefore nothing improper about the idea that for convenience any actual (e.g. English) words can figure in text representations, as long as they do have formulae. Thus one might elaborate

a text representation with a rather particular word like “aeroplane”, apparently far more specific in meaning than, say, THING, as long as “aeroplane” has its own THINGy formula. Yorick argues that there is no more circularity about this than about any dictionary, and notes that a detailed study of Webster’s dictionary showed some rather generic words recurring in many definitions, functioning like primitives. This position is different from the Katz and Fodor one with markers subsuming distinguishers.

Wilks (1977a) is a more substantial discussion of the pros and cons of semantic primitives, or rather of what Yorick believes is their essential character and hence rational justification, contra others including Katz and Fodor, and Putnam, for example. Thus he offers criteria for (a set of) primitives: being finite, comprehensive, independent, non-circular and non-reducible. Further, a set of primitives (with an application syntax) is a reduction device yielding a semantic representation for a natural language via a translation algorithm which is not plausibly explicated by other entities of the same type (p.184). But aren’t they really things of another type? Mental forms, for example? How could one know? Or terms in a model-theoretic formal language, Markerese? But model theorising is not what natural language processing is all about. Guaranteeing, say, that “seek” (or ‘seek1’) is equivalent to ‘TRY TO FIND’ is a lost cause for natural language. Yorick’s position is essentially that trying to capture natural language semantics via meaning postulates is trying to bag a black cat: when you open the bag there is only a Cheshire grin. Semantic primitives are nothing to do with analyticity, or with stereotypical facts: the former is too restrictive and the latter too demanding for a necessarily reduced representation language.

Primitives constitute a language for the description of meaning (p.191) and are thus, as mentioned earlier, no more univocal than any other language. Yorick’s position is that (both non-case and case) primitives are a useful organising principle for a natural language processing system because they allow helpful generalisations to be made, as operationally convenient: he endorses Sampson’s analogy of primitives as like English pound notes with promises to pay that mean the notes can be turned into something else but never into actual gold. Just as there is no gold that English pound notes can be turned into, there is no conceptual substance that primitives can be turned into. As Goodman had earlier claimed, a primitive representation language is just that, a language, among many others, with no absolute status. Of course the machine has to know the language, or rather we as its program writers do, but this is no different from dealing with any other language, formal or informal. Yorick further argues that when studies of large conventional dictionaries show that some words feature conspicuously as recurrent defining terms for other words, they are organisational devices because, though they themselves have dictionary definitions, these definitions are mere empty gestures, not substantive ones.

It is worth noting that where I have just referred to language *processing* systems, Yorick during the 1970s referred to language understanding systems: this is somewhat at variance with his general stance on primitives, and is perhaps attributable to the fashionable terminology of the period, or at any rate should be interpreted as ‘understanding sufficiently for the task in hand (e.g. translation)’.

The ultimate justification for some language of primitives is thus whether it works for some language processing purpose, on some suitable test of working. In Wilks (1977b) and (1978), Yorick further extended his apparatus to deal with two problems of practical language processing, and in so doing further amplifies the points just made.

In (1977b) he grapples with the need to exploit different underlying causal-type relations, and specifically to distinguish between ‘cause-of’ and ‘reasons-for’, to link different parts

of a text. He maintains that this can be done, within the overall preference framework, by classifying inference rules so that, essentially, they operate in one semantic direction for CAUSE and the other for GOAL. The result is rather complicated and it is far from clear that what is illustrated for a few examples will scale up satisfactorily.

Pseudo-texts, and metaphors

Wilks (1978) is a more direct development of the line Yorick took in Wilks (1977a) because it enriches the language-like characteristics of his approach, and of the mechanisms it supports, to deal with the key fact of ordinary language use, namely the continuous appearance of new word senses. It is also Yorick's response to the contemporary interest in organised bodies of knowledge about the world (frames etc.). Thus Yorick's proposal is to import another class of primitive-based resource, *pseudo-texts*, that are frame-like objects that encapsulate detailed world knowledge and can be invoked to enlarge any existing text representation using templates, paraplates etc. and, specifically, make possible the connective inferences that provide an interpretation for new word senses. The particular point of interest is the way that Yorick's existing apparatus, relying wholly on primitives, can be connected with pseudo-texts, which may need to be much more specific: many words for types of weapon will all have the same formula, for example, but a particular text may require some more specific characterisation of guns to be understood which the "gun" pseudo-text supplies.

Yorick's way of bridging this gap is to exploit the thesaurus idea. Thus he notes that primitive formulae may be taken as imposing a structure on a thesaurus in that the individual primitives within a formula context can point to word classes in a thesaurus. This imposed structure is richer than the normal simple hierarchy in a thesaurus like Roget's (just as Yorick's formulae are themselves more complex than sets of class labels for word senses). However the thesaurus class hierarchy is also exploited to enhance the formula-based apparatus. Thus the classes pointed to may be whole head classes, or subheads, or fine-grained bottom level synonymous word sense sets, or *rows*. Any individual word (sense) can be invoked as long as it has a formula, and so implicitly substituted in a template etc. Word classes can also be invoked since there are formulae, with inclusion relations, for the word sets at different levels within a thesaurus head, bottoming out at row level. (It is not clear whether these formulae are inferred from shared parts of bottom-level formulae or are specifically constructed.)

The pseudo-texts embodying word-related facts associated with particular words or word classes are the same sort of primitive-based pattern structures as ordinary text representations, so when they are invoked through the words or word classes that can fill slots in the initial text representation, they provide more template, etc. patterns across which inferences can be made. The presumption is that all this invocation will occur only when existing representation patterns clash with the available preferences. The operation is envisaged as an inferential projection from the pseudo-text onto the actual text representation, supplying new patterns which are near enough those sought to do as interpretations: thus if we cannot interpret "drink" in "My car drinks gasoline" directly, we can get a good enough interpretation by importing "use" from the pseudo text for "car".

As this suggests, Yorick's apparatus is becoming increasingly baroque. He admits that there is no implementation, and that there are important issues like control to tackle. His claim is that his approach combines coherence in the use of the same kind of primitive-based representation across many different components with flexibility through need-driven and

preference-based use of these components, but the evidence that this strategy actually works is missing.

Yorick's last major paper on the line just laid out was Fass and Wilks (1983). Much of the apparatus, and illustrated application, is as before, with a discussion of "My car drinks gasoline" as a metaphor and comparison between the way it is treated in Preference Semantics and in other approaches to metaphor. However there are some differences. Thus in considering how to handle metaphor as a normal feature of language use, the paper again argues for the relative preferences approach as a better way of accommodating 'ill-formed' input than fixed semantic constraints. But the paper flags a more explicit emphasis on the forms of semantic information used as types of *dictionary* information. Thus it argues that dealing with metaphors is better done by allowing some weakening (by generalisation or partial matching) of either the core elements of dictionary definitions, or their relational conditions, than by invoking the frame-like pseudo-texts referring to world knowledge of Wilks (1978). The notion is that the character of the metaphor is displayed by the nature of the modification.

Preference Semantics' claims

Fass and Wilks (1983) also marks another shift, and is of particular significance here, apart from its concern with metaphor, for two reasons.

First, it states, quite specifically, that Preference Semantics is not a set of programs, but a set of principles or claims. It is thus a system only in an abstract sense, to be justified less by actual implementational achievements, than by the merit of its general claims and by concrete illustrations of the way it works. Yorick comments, more than once (e.g. in Wilks 1975c), that there was an implementation, with a 600-word dictionary, that did process a range of English paragraphs successfully, by translating them into French. There is thus more support for his illustrations than what might be called the usual theoretical linguists' style of illustration. But there is no commitment to making a serious application system work, come what may.

Second, Fass and Wilks lists the claims that Preference Semantics makes: summarising, these are

1. there is no syntactic module;
2. the semantics are not model theoretic, and quantification just needs some special procedures;
3. everything is procedural, and generally so, operating under a least-effort principle;
4. there is *some* privileged set of semantic primitives;
5. text representation is linear and primarily surface-text sequenced;
6. the representation of a text is the best fit among competitors;
7. hence ill-formedness is only relative not absolute, and does not preclude interpretation.

There is some irony in the fact that these claims appear at what was in fact the end of Yorick's work on Preference Semantics as we knew it, given not only that some claims are not supported by much evidence, e.g. that on quantification, and that, as we shall see, some concurrent attempts to exploit Preference Semantics ran into difficulties. But as against

this, a fairer view is that Yorick’s move in the 1980s to concentrate on automating lexicon construction was the correct strategy: even without a commitment to building lexicons for practical natural language processing systems and to cost-effective methods of doing this, it is necessary, for intellectual credibility, to show that it is possible to build a non-trivial lexicon so that a lexically-based approach to text interpretation can be independently tested.

From this point of view, Yorick’s move was *reculer pour mieux sauter*. But Yorick’s 1965-1985 version of Preference Semantics then vanished from sight. I will return to its reappearance, or reincarnation, later. Its retirement, at least into the wings, in the mid 1980s is not wholly surprising: quite apart from the scaling up challenge, and the fundamental question of whether Preference Semantics was indeed the right way to approach semantic interpretation, there were other factors in play. One was the enthusiasm, from the late 1970s through the 1980s, for the logico-grammaticist approach to language processing and meaning representation, with its emphasis on logical form, its more limited interpretation of “semantic”, and its sharper distinction between semantic, in this sense, and pragmatic. Thus there was both more concern with issues like the treatment of quantifier structures, which Yorick in practice ignored, and a harder line about what constitutes the genuinely linguistic information about word meanings to be embodied in dictionary, i.e, system *lexicon*, entries.

Yorick’s apparatus indeed depended on the lexical formulae for words, but also on the other pattern sets. In the more lexicalist approach of the 1980s, the narrowly linguistic share of the pattern information tended to be dispersed in the form of constraints on individual words contained in their lexicon entries. These entries, expressed in feature set form, indeed made use of general semantic categories and were typically supported by subsumption hierarchies, and so implicitly involved general patterns. But even if there were elements of common linguistic description behind such thoroughly different ‘notations’, there was also a much more important and radical difference between Yorick’s approach and the then dominant ones in the key role of syntax. Processing within the grammaticist approach was dependent on and driven from syntax, even if the eventual meaning representations were not conventional syntactic parse trees. The grammatico-logicist approach also led to ‘deep’ representations with structures rather far removed from the surface text. As Yorick commented in Wilks (1983b), ‘deep’ and ‘superficial’ (aka surface) are complicated notions. In his case the primitive formulae are deep but his text representations, with their sequences of templates, are actually shallow.

4 Applications

These competing developments were reinforced by the attempts others made to exploit Yorick’s approach. His own experiments and implementations were not so compelling as to lead many others with natural language processing interests to follow him, or to supply exportable technology; but it is worth commenting on Boguraev’s work, first on ambiguity resolution per se and then on the database query task, as applications and developments of Yorick’s key ideas (Boguraev 1979, Boguraev and Spärck Jones 1983).

The main features of Boguraev’s initial work were a return to conventional syntactic parsing as an essential component of processing, and the derivation of a semantic representation as a case-labelled dependency tree over what would nowadays be called predicate-argument or proposition-like constituents. Boguraev’s work was with individual sentences, rather than extended texts. His system used ATN parsing that exploited dictionary entries combining

conventional syntactic information with Wilks-type semantic formulae. His concern, like Yorick's, was with both lexical and structural ambiguity (especially associated with prepositions). Processing was driven off contextual verb frames, with additional semantic patterns much like paraplates, called *preplates*, for dealing with modifier attachment. Over time Boguraev's work came to involve a larger set of case labels than Wilks', which could be supplied from prepositional lexical entries or in other ways. Boguraev's basic mechanism was, however, the same as Yorick's, namely the use of semantic preferences.

Boguraev's semantic sentence representations were explored for a more autonomous generation process than Wilks and Herskovits' translation into French: thus he demonstrated successful ambiguity resolution via paraphrase. At the same time, in the application to database query, the representation could also be used as input for further transformations of an English question into a formal database query. In particular the detailed dependency tree could be used to handle quantified structures in the precise way required by database query, but very doubtfully possible with Yorick's original system. For both the ways of exploiting his semantic representations that Boguraev studied, the dependency tree structure provided significant leverage.

This line of work, which appears to be the most substantial attempt to apply Yorick's ideas outside his own group, eventually petered out, partly for external reasons, but partly for internal ones. The database frontend work, like other such work elsewhere, came up against the challenge of supplying domain models and the need for robustness in connecting human users with hidden formal data structures. However, as with Yorick's own work, the business of enlarging the lexicon in an adequate and consistent way led naturally to Boguraev's own work on ways of exploiting machine-readable conventional dictionaries as sources for processing system lexicons. The difficulties of supplying Wilks-style lexicon entries had also led, in another project on text-retrieval requests (Spärck Jones and Tait 1984), to a radical simplification of the formulae, making their semantic content and their relation to syntactic information closer to the conventional one.

5 Building machine lexicons

Yorick's work on automated lexicon construction was more than a simple response to the data and data processing capabilities that were becoming available at the time. Exploiting existing machine-readable dictionaries as the base for new ones designed for natural language processing systems, or exploiting corpora to extract word behaviour data, is entirely in the spirit of his long-standing views about the nature of his approach to semantic interpretation. But it also marked at least an apparent change in his detailed approach to semantics.

This is fully evident in Wilks et al. (1996). *Electric words* also provides a retrospective overview, in the context of theories of meaning, of the general (though not specific) approach to semantic representation that Yorick adopted in the work I have described. This approach is offered both as a sound approach to linguistic meaning representation in its own right, and as an appropriate basis for the strategy of building language processing lexicons by bootstrapping from machine-readable dictionaries (MRDs) that is now proposed, to meet the scaling up challenge, as the way forward for the field. The message in all of this is, again,

- 1) that word meanings can (in general) only be conveyed by other words, i.e. through some other language which necessarily has the properties of any natural language, like lexical ambiguity;

2) that semantic primitives provide anchoring pegs, or an organising apparatus, for this semantic characterisation;

3) that any set of semantic primitives is the right set only because they work, as a specialised sublanguage, in enabling some language processing task, like translation;

4) that such useful primitives emerge, as their motivating ontology, when an existing dictionary is analysed as a text.

These are very general statements. The point of interest here is precisely what form these semantic primitive-based entries in the language processing lexicon derived from a conventional dictionary are like, along with what this implies for the the other system contributors to text processing and for the form of text representation these deliver, say for translation. In fact, of course, the language processing lexicon will not simply rise out of the MRD like Venus from the waves: the process will be more like fishing with some carefully chosen bait; and it will also gain, taking it yet further from the distributional purist's approach to pulling all linguistic units and structures out of running text, by starting from the 'preprocessed' text corpus that an existing dictionary text in itself provides.

This is not the place to recapitulate the detail of Yorick's group's work with MRDs: I want only to consider its key points, especially as illustrated in Wilks et al. (1987) and in *Electric words*, and how these consort with Yorick's earlier Preference Semantic system.

Wilks et al. (1987) describes several independently-pursued lines of work, with LDOCE (Procter 1978), but these can all be related to different aspects of Preference Semantics. First, and most important, semantic primitives survive. But there are far more of them, around 1000 terms identified (through frequency and simplicity) as central in the 2000+ terms of the basic vocabulary used for LDOCE definitions. This is far more than Yorick's original set of less than 100, though the suggestion is that the 1000 can be further reduced. The claim, on the basis of experiment, in Wilks et al (1987) is that provided these terms themselves are properly, de facto manually, defined, the definitions for the much larger word set in LDOCE can be automatically rewritten, in bootstrapping cycles, to obtain a derived dictionary with primitive-based definitions.

This in turn provides the material for machine lexicon entries which are in frame form, and can in principle be extracted by automatically parsing the dictionary definitions. These frames combine both linguistic knowledge and world knowledge, and also encode case relationships, as well as conventional grammatical information. These frame structures, again experimentally investigated, combine the types of information that in the earlier Preference Semantics apparatus was spread across formulae, templates and paraplates, and pseudo-texts. However the primitives function much more as simple category labels, and the distinctive syntax of Yorick's original word sense formulae appears to have vanished.

At the same time the core lexical data could be enhanced in two ways. First, by processing dictionary entries to extract genus hierarchies which would allow generalisation and inference in the way that classes of primitive did in Yorick's original Preference Semantics; and second by processing the dictionary as a corpus to extract word cooccurrence relations and, in principle, classes of words with similar dictionary-text behaviour: these would function in the same way as rows in Yorick's earlier system, as links between individual words with their own distinctive properties and the primitive formula characterisations of word sense classes.¹

¹Yorick comments on the relation between this work and some of my earlier work, but I am ignoring this here.

Similar ideas figured in other research on MRDs in the late 1980s. But quite apart from the political intellectual property obstructions this work ran into, there were far more substantial barriers to progress. The most salient was correct word sense identification in the dictionary entries, i.e. sufficiently reliable sense selection to drive the whole bootstrapping process. But there was also, especially for the present context, the awkward fact that a ‘good’ (even if not perfect) set of semantic primitives has not been found to emerge. Yorick’s view that you cut your language processing primitive cloth for your application purpose suit is a hard dialectic taskmaster.

Current approaches to the language processing lexicon illustrate different responses to these problems, but collectively diverge from Yorick’s primitive-based lexical centre. In one strategy, as implemented for the LKB (LKB 2005), lexical entries have a rigorous formal (typed feature) structure, but have only very general semantic category features, like ANI, and relational features, like ‘telic’, alongside syntactic data. In a complementary approach, as manifest in WordNet (WN 2005), there is a rich, fine-grained descriptive word classification, in the same ball park as Roget’s, enhanced with some simple syntactic category and verb frame information. Both strategies have more limited aims than the earlier ones, but have more chance of being able to automate at least some of the lexicon building work, as with EuroWordNet (EWN 2005). FrameNet (FN 2005) is somewhat closer in spirit to the kind of lexicon envisaged in, or rather implied by, Wilks et al (1987). It offers both conventional syntactic data and row/synset word sense class data, thus (roughly) combining LKB and WordNet-style approaches. But in addition, and most importantly, is organised by semantic frames, like Activity or Ingestion, that are effectively a large number of low-level semantic category primitives, and by a set of semantic case primitives, like Manner or Place, that define frame slots. However FrameNet also conveys much of its lexical information by the most straightforward use of ordinary, detailed English without any concession to limited defining vocabularies, e.g. “Means of Ingestion: an act performed by the Ingestor that enables them to accomplish the whole act of ingestion”. The detail involved also implies that, as with conventional lexicography, building FrameNet is a primarily human activity.

A direct semantic comparison between an early Wilks lexical entry and a FrameNet one is hardly fair, because Yorick relied on explicit, separate pattern sets, notably the paraplates, to supplement individual word information, where Framenet supplies this information directly in its lexical entries. But even so, just considering Yorick’s entry for “grasp”, in the sense of grasp an idea, and the corresponding word sense one in FrameNet, as given in Figure 1, show how different Yorick’s original and one major modern view about the nature and role of primitives in a general-purpose lexicon are. What is less clear is how different the outcome of Yorick’s ideas about the lexicon in what we may call his MRD-based phase, as illustrated by Wilks et al. (1987), and this modern product would be.

Nor, of course, do we have any idea about how effective any of these lexicons would be for real, tough natural language processing tasks. A great deal of use is made of WordNet, but much of this is because it is all there is and it has been found to be of some use. Noone would say this implies it is the optimal semantic lexical resource. The very limited and most general semantic categories, like ANI or PHYSOBJ recur in many dictionaries, with the same motivation as Yorick’s formula head primitives. The question is whether there is value in additional modifying category primitives (not case ones) as well for semantic processing: grasp is not merely THINK, but a recognising TRUE BE sort of THINK. Again, corpus-based lexical sets and relations have been found practically useful. But this cannot be taken to demonstrate that all the linguistic reality that semantic primitives in some strong sense do

appear to embody is wholly captured, even if left implicit rather than explicitly labelled, by current statistical operations.

Wilks ‘grasp’ FORMULA :

(ANI SUBJ)(SIGN OBJE)((THIS (MAN PART)) INST) (((SAME SIGN)(TRUE BE)) THINK)

FrameNet ‘grasp’ [syntax detail], FRAME :

GRASP

Definition:

A Cogniser possesses knowledge about the working, significance, or meaning of an idea or object, which we call Phenomenon, and is able to make predictions about the behaviour or occurrence of the Phenomenon. The Phenomenon may be incorporated into the wider knowledge structure via categorisation, which can be indicated by the mention of a Category. The Cogniser

Frame Elements:

Core:

Cogniser The sentient animate being who acquires new knowledge.
Semantic Type-
 Sentient

Faculty A part of a person’s cognitive-emotional faculties that is said to acquire knowledge.

Phenomenon A state of affairs or dynamic system whose internal makeup and working the Cogniser comes to assimilate into their knowledge structure.

Non-core:

Category This expresses a general type or class of which the Phenomenon is considered an instance by the Cogniser, allowing them to make predictions about the qualities, occurrence or behaviour of the Phenomenon.

Completeness The extent to which the Cogniser has incorporated the workings and significance of s Phenomenon into their knowledge structure.

Evidence ...

Manner

Reference-point ...

Time ...

Figure 1

6 The iceberg

Returning now to my starting point, and the iceberg with which I began: the review of theories of meaning in Wilks et al. (1996) reminds us that the underwater part, below the semantic primitives tip, is very large and, spreading wide as well as deep, often very deadly. Much natural language processing (though not computational linguistic) work is pursued with a purely practical attitude, on the ‘go for it and if it works, fine’ principle. Who cares about the underwater berg of theories of meaning if, as long as you are careful about data detail, you can get your system to work? But as Yorick’s papers over the long period considered here show, you cannot build practical language processing systems without adopting some position about the base on which you are building, and foundations are not only theories, they are even metaphysics. This applies just as much to currently fashionable statistical approaches, where language models are only another form of dictionary and computing mutual information is only an unconventional form of parsing. These statistical approaches are appealingly austere and apparently without any metaphysical baggage. They nevertheless rest on a theory about how meaning is recognised, represented and manipulated, just like other more obviously theory-laden approaches do. Here Yorick’s ‘language’ account of meaning shows the iceberg with a less dangerous bulk than some other theories have.

But theories of meaning in general are not all there is to the underwater iceberg. There is also the closely related question of linguistic creativity, and notably of metaphor, where the kind of interpretive strategy that Yorick advocated is one way of not getting wrecked. Again, though Yorick’s position is implicit rather than explicit, his use and view of primitives offers one account of language universals. Since his semantic primitives started out as tools for conventional, bilingual translation they must, if effective, have some degree of universality. More generally, his account of meaning representation through the use of a meaning representation language which, however limited it is, has some crucial natural language properties like word sense ambiguity, implies *some* universality property for his primitives. But this is an Aristotelian, not a Platonic, account of universals.

Equally, Yorick’s view of primitives and its grounding in the larger contextual-procedural approach to meaning determination sits on top of a whole iceberg mass of theories about language and the form and role of computational linguistics, as opposed to applications-oriented language processing. Here Yorick’s emphasis on language process and action rather than simply description is entirely right, though there is no evidence that mainstream linguistics has taken any notice of this defining methodological and substantive contribution from computational linguistics to linguistics in general. However, though Yorick has emphasised the claim that his approach embodies syntactic as well as semantic conditions, so one does not need a distinct syntactic processor, one does not have to be a formal semanticist obsessed with quantification to feel that there is more to the part of the iceberg to do with the modules and architecture of an abstract language processor than this; and this applies both to the computational case and the cognitive one. Here modern approaches, where some (though perhaps weak) form of semantic primitive provides the bridge between lexical, syntactic and semantic components, of this kind illustrated earlier, are more convincing.

One other chunk of the iceberg below Yorick’s earlier work also deserves comment. His approach to disambiguation relied not merely on the wider text context but on this larger context having a particular form of representation: his form of discourse representation was a shallow one, ordered and hence parallel with the surface text. This puts it in the same general class as Rhetorical Structure Theory as opposed to Schankian scripts, say, but the

underwater iceberg here is vast. From this point of view it is a pity that Yorick’s move to machine lexicons diverted him from what became a major area of research.

7 Primitive preferences: where are we now?

In the final review, what does Yorick’s use of semantic primitives offer us? The crux here is the one I touched on earlier, in discussing machine lexicons drawn from MRDs, namely how similar Yorick’s and, insofar as they use semantic primitives, modern meaning representation *languages* actually are. This is not just a matter of whether their primitives have the same names or whether, when statistical derivation is in question, one can find primitive names: as against Yorick’s claim in Wilks et al. (1996) that one cannot, one can point to obvious strategies like taking the most frequent *word* from a set of grouped rows. In such a case one might get “act” – > ACT, but this is near enough Yorick’s DO, especially when on Yorick’s own principles one does not suppose that there is any one intrinsically correct primitive set. One of the engaging features of Richens’ original NUDE, and of Yorick’s development of it, was just how natural it was as a paraphrasing language, and even how joyous it was to use as a language of communication: what is a surprise but a BANG DO ? One might imagine re-forming modern feature-structured lexical entries to give them a Wilksian shape, but the result is far from a stimulating pidgin: developing the illustration for “ammeter” and “measure” in Wilks et al. (1996) we might get (measure PURPOSE)(Solid/Movable PHYSOBJ) and (Human/Sex-Unspecified SUBJE) (Abstract OBJE) ACTION which, though perfectly reasonable, lack zap.

Of course this is not exactly what these things are for. WordNet, FrameNet, MindNet (MN 2005), the annotation schemes using in such banks as PropBank (PB 2005) are all present players in the semantic lexicon space, and WordNet in particular is applied in practical task systems, for example for question answering. But as mentioned earlier, WordNet has been applied primarily because it is available, in system-usable form. It, and these other resources have been built as general-purpose resources (as Roget’s original *Thesaurus* was), and have not always been of value for particular tasks. They have been seen more as descriptive than as task and process-oriented, but though descriptions they can still be exploited predictively when amplified with appropriate application rules. There is thus, I maintain, no generic difference between them and Yorick’s apparatus as primitive-exploiting semantic tools, however large the detailed differences are. There are real questions, currently being rerun in the Semantic Web and ontologies world, about the relations between, and values of, domain-independent and domain-dependent semantic structures and about the relations between linguistic and world knowledge. There can be no unequivocal answers to these questions, as Yorick always recognised: his position was that while you can’t make a language processor without semantic primitives somewhere, you choose your semantic primitive cloth, and tailor it, to suit your processor climate.

So we should bear in mind, as possibly suggestive for the future, Yorick’s Preference Semantics approach to text interpretation and the reasons he advanced for it, perhaps acknowledging his message thus:

(POINT : MAN [yw]) / (CAUSE / (FOLK / ((THINK : SIGN) : FEEL)))

References

This list of references has two parts. The first contains items by other authors than Wilks that are cited in the text. The second contains items by Wilks that are cited in the text, in their temporal order. This second part is not exhaustive, but consists only of those items relevant to, and discussed in, the text.

Boguraev, B.K. (1979) Automatic Resolution of Linguistic Ambiguities. PhD Thesis, University of Cambridge, 1979; Technical Report 11, Computer Laboratory, University of Cambridge, 1979.

Boguraev, B.K. and Spärck Jones, K. (1983) How to Drive a Database Front End Using General Semantics. Proceedings of the Conference on Applied Language Analysis, Association for Computational Linguistics, 1983, 81-88.

EWN (2005): EuroWordNet, see <http://www.illc.uva.nl/EuroWordNet/> (visited 2005).

Fillmore, C.J. (1968) The Case for Case. In Bach, E. and Harms, R. (eds.) Universals in Linguistic Theory. New York: Holt, Rinehart and Winston, 1968, 1-88.

FN (2005): FrameNet, see <http://framenet.icsi.berkeley.edu/> (visited October 2005).

Herskovits, A. (1973) The Generation of French from a Semantic Representation. Memo STAN-CS-73-212, Stanford Artificial Intelligence Laboratory, Stanford University, 1973.

LKB (2005): LKB, see <http://wiki.delph-in.net/moin/LkbTop> (visited October 2005)

Masterman, M. (2005) Language, Cohesion and Form. (ed. Y. Wilks), Cambridge: Cambridge University Press, 2005.

MN (2005): MindNet, see <http://research.microsoft.com/nlp/Projects/MindNet.aspx> (visited October 2005).

PB (2005): PropBank, see <http://www.cis.upenn.edu/ace> (visited October 2005).

Procter, P. (1978) (ed.) Longman Dictionary of Contemporary English. Harlow, Essex, UK: Longman.

Spärck Jones, K. and Tait, J.I. (1984) Automatic Search Term Variant Generation. Journal of Documentation, 40, 1984, 55-66

Spärck Jones, K. (2000) R.H. Richens: Translation in the NUDE. In Hutchins, W.J. (ed.) Early Years in Machine Translation. Amsterdam: John Benjamins, 2000, 263-278.

WN (2005): WordNet, see <http://wordnet.princeton.edu> (visited October 2005).

Wilks, Y. (1965) Computable Semantic Derivations. ML 176, Cambridge Language Research Unit, 1965 (see also Computable Semantic Derivations, Report SP-3017, System Development Corporation, Santa Monica CA, 1968, and Chapter 3 in Wilks (1972)).

Wilks, Y. (1967) Semantic Consistency in Text - an Experiment. Report SP-2758/000/00, System Development Corporation, Santa Monica CA, 1967.

Wilks, Y. (1971) Decidability and Natural Language. Mind, LXXX, 497-520.

Wilks, Y. (1972a) Grammar, Meaning and the Machine Analysis of Language. London and Boston: Routledge, 1972.

Wilks, Y. (1972b) Lakoff on Linguistics and Natural Logic. Memo STAN-CS-73-457, Stanford Artificial Intelligence Laboratory, Stanford University, 1972.

Wilks, Y., and Herskovits, A. (1973) An Intelligent Analyser and Generator for Natural Language. Proceedings of the International Conference on Computational Linguistics, Pisa,

1973, (ed. A. Zampolli). Florence: Olschki, [?1974 KSJ's copy different biblio detail]

Wilks, Y., and Schank, R. (1974) The Goals of Linguistic Theory Revisited. *Lingua*, 34, 1974, —. (originally Memo STAN-CS-73-368, Stanford Artificial Intelligence Laboratory, Stanford University, 1973.)

Wilks, Y. (1975a) Preference Semantics. In Keenan, E.L. (ed.) *Formal Semantics of Natural Language*. Cambridge: Cambridge University Press, 1975, 329-348. (originally Memo STAN-CS-73-377, Stanford Artificial Intelligence Laboratory, Stanford University, 1973.)

Wilks, Y. (1975b) An Intelligent Analyzer and Understander of English. *Communications of the ACM*, 18 (2), 1975, 264-274.

Wilks, Y. (1975c) A Preferential, Pattern Seeking, Semantics for Natural Language Inference. *Artificial Intelligence*, 6, 1975-53-74.

Wilks, Y. (1975d) Primitives and Words. *Proceedings of the Workshop on Theoretical Issues in Natural Language Processing (Tinlap)*, (ed. R. Schank and B. Nash-Webber). Association for Computational Linguistics, 1975, 38-41.

Wilks, Y. (1977a) Good and Bad Arguments About Semantic Primitives. *Communication and Cognition*, 10 (3/4), 181-221.

Wilks, Y. (1977b) What Sort of Taxonomy of Causation Do We Need for Language Understanding? *Cognitive Science*, 1, 1977, 235-264.

Wilks, Y. (1978) Making Preferences More Active. *Artificial Intelligence*, 11, 1978, 197-223, and in Findler, N.V. (ed.) *Associative Networks*. New York: Academic Press, 1979, 239-266.

Wilks, Y. (1982) Some Thoughts on Procedural Semantics. In Lehnert, W.G. and Ringle, M.D. (eds.) *Strategies for Natural Language Processing*. Norwood NJ: Lawrence Erlbaum Associates, 1988, 495-516.

Wilks, Y. (1983a) Does Anyone Really Still Believe This Sort Of Thing? In Sparck Jones, K. and Wilks, Y. (eds.) *Automatic Natural Language Parsing*. Chichester: Ellis Horwood, 1983, 182-189.

Wilks, Y. (1983b) Deep and Superficial Parsing. In King, M. (ed.) *Parsing Natural Language*, London: Academic Press, 1983, 216-246. Reprinted in Woods, W.A. and Fallside, F. (eds.) *Computer Speech Processing*. New York: Prentice Hall, 1985, 335-362.

Fass, D. and Wilks, Y. (1983) Preference Semantics, Ill-Formedness, and Metaphor. *American Journal of Computational Linguistics*, 9 (3/4), 1983, 178-187.

Wilks, Y. et al. (1987) A Tractable Machine Dictionary as a Resource for Computational Semantics. Memo MCCS-87-105, Computing Research Laboratory, New Mexico State University, Las Cruces, 1987.

Wilks, Y. et al. (1989) A Tractable Machine Dictionary as a Resource for Computational Semantics. In Boguraev, B. and Briscoe, T. (eds.) *Computational Lexicography for Natural Language Processing*. London: Longman, 1989, 193-228.

Wilks, Y.A., Slator, B.M. and Guthrie, L.M. (1996) *Electric Words*. Cambridge MA: MIT Press, 1996.