MODULARITY AND LINGUISTIC AMBIGUITY

Robyn CARSTON

Abstract

Jerry Fodor has proposed that language processing is modular, that
is, that the processes responsible for determining the grammatical
and logical structure of an utterance do not (and cannot) use
background or contextual information even when it would be relevant.
They are informationally encapsulated. This thesis has provoked much
reaction, especially from Marslen-Wilson and Tyler who interpret much
of their experimental work on utterance comprehension as providing
evidence against the modularity of language processing. In this
paper I consider some of the issues at stake and look particularly at
Marslen-Wilson and Tyler's strongest piece of putative counter-
evidence, involving the parsing of ambiguous syntactic structures
such as 'shaking hands'. I show that their analysis depends on a
failure to distinguish facts about syntactic structure from facts
about pragmatic plausibility, and that once this distinction is made
the modularity position remains untouched.

1. **Background**

        A central issue in the study of cognitive systems is the extent
to which specific domains (language, vision, auditory perception)
function autonomously, in accordance with their own specific
principles and proprietary data base, and the extent to which they
call on principles common across domains and information from a
range of sources. There is a continuum of views possible, from a
strictly bottom-up modular view to a strongly interactive position.
A modular view of mental processes along the lines of Fodor (1983,
1986) allows for module-internal interaction and for a weak
interaction (to be explained below) between the modular processor and
the central systems domain of conceptual/propositional represent-
ations. In his view modular systems tend to have a cluster of
properties which distinguish them from non-modular systems: i.e. they
tend to be specialised computational systems, fast and mandatory in
their operations, hard-wired, innately specified, and, most
importantly in his view, informationally insulated: "Computational
modules are domain-specific and they're also computationally
(informationally) encapsulated in the sense that not all the
information available to the organism for some task or other needs
to be available to any given module for the performance of its
computations." (Fodor, 1986, 12).
        Among the mental processes that he proffers as forming such a
module are those of the language faculty, in its function as an input
system, i.e. a system which converts transduced linguistic stimuli

into representations which can interact inferentially with other higher level, presumably conceptual, representations which may have been derived from other sensory stimuli, or be longer-standing assumptions, retrieved from encyclopaedic memory.

This view of language processing as modular should be distinguished from others in the literature, as it is far less rigid than many. It doesn't entail strict seriality of processing, in particular it doesn't entail the clausal hypothesis, whereby the unit of syntactic analysis is taken to be a clause and any interaction with semantic context is precluded until the clause boundary is reached. According to Fodor "... semantic information is never used to predict syntactic structure, but a line of analysis on which the parser is engaged can be aborted whenever it produces structures that resist contextual integration." (Fodor, 1983, 134-5). That is, contextual information comes in to assess the analysis bit by bit as it is computed by the language system; feedback is limited to acceptance and rejection. There is a real question as to what the "bits" or units of analysis are, but at least it is evident that Fodor imposes no a priori restrictions on this, such as a whole clause. His is very much an on-line interactive theory, though with certain restrictions on the nature of the interaction.

Secondly, while precluding top-down information flow from central conceptual sources Fodor doesn't preclude top-down flow internal to the module, i.e. he's happy to explain such well-attested phenomena as phoneme restoration and the word superiority effect in terms of higher levels of the language module affecting lower levels. (These background issues and others are discussed in greater detail in Carston 1988.)

2. Word recognition

A model of word recognition which seems to me to perfectly exemplify the picture given above is the revised cohort model of Marslen-Wilson (1987). Before briefly considering this though, a word or two on the original cohort model (see Marslen-Wilson and Tyler, 1981). If we recognise three phases in word recognition: access, selection and integration, the access phase is entirely bottom-up (stimulus-driven), while the selection phase, in the first version of the cohort model, involves an interaction between the dictates of the sensory stimulus and contextual constraints imposed by the place that the word is to take in the evolving syntactic structure and in the higher level conceptual representation. So, roughly, the first couple of phonemes activate all of those words in the listener's mental lexicon which begin with that initial sensory sequence, making up what is known as the word-initial cohort. Then, if the word is part of an utterance, rather than being heard in isolation, syntactic and interpretative specifications, along with the continuing sensory input, cause candidates to drop out of the pool. Selection has occurred when a single word candidate remains which matches both the context and the sensory input.

This highly interactive picture has always seemed to me to be somewhat at odds with the results of experimental work by Swinney

and his colleagues, despite Marslen-Wilson and Tyler's approving reference to them (Marslen-Wilson and Tyler, 1981, 114). What these results show is that upon encountering an ambiguous word, a homophone, hearers can't help but select both lexical items conforming to the phonological type, even when the overall context is clearly biased in favour of one sense over the other. An instance of the crucial kind of example is:

(1) The man was not surprised when he found several spiders, roaches and other bugs$_\triangle$ in the corner of his room.

Subjects were presented auditorily with such sentences and were required to make a lexical decision regarding a string of letters presented to them visually at the acoustic offset of the ambiguous word, marked here by $\triangle$. Reaction times to words conceptually or associatively related to either of the senses of the homophone (such as 'spy' and 'ant' in this example) were found to be equivalent and significantly faster than reactions to words unrelated to either sense (such as 'sew'). So words related to either sense are primed and the preceding context (which in this example strongly biases for the insect sense of 'bug') has no effect on the selection process. That is, the two competing candidates remain active at the point when the complete sensory stimulus has been received, contrary to the predictions of the first cohort theory. It should be noted that the homophones tested included a large number of two syllable cases ('panel', 'organ', 'table', 'duty') for which the same result arose. This is important since it might be possible to argue that with at least some one syllable words, such as 'bug', there isn't time for contextual constraints to exercise any influence before the entire sensory input is received and completes the job.

The results seem to be particularly robust, having been replicated several times by Swinney and his colleagues and by Tanenhaus and his (see Seidenberg, Tanenhaus, Leiman and Bienkowski, 1982, and, for a general overview of relevant work, Tanenhaus and Lucas, 1987). Even when there are large differences in frequency of one of the two senses, as is the case with 'pen' and 'scale' and a contextual bias is set up which favours the sense which is also the more frequent one, even then, words related to both senses are primed and to roughly the same magnitude, (see Onifer & Swinney 1981).

Of course context does very rapidly have an effect, within 3 syllables on Swinney's account, and perhaps even earlier, so that when visual presentations were made at /corn $\triangle$ ar/ only 'ant' was facilitated: i.e. a lexical decision for it was made significantly more quickly than for 'spy' which fell in with 'sew' at this later point in processing. At this (post-modular) stage, integration of the word into the higher level representation has taken place, with the rejection of the other sense. This is not at all surprising and doesn't bear on the issue of whether processes of word recognition (in particular selection) are modular or not.

Let's look at the predictions of the first cohort model, taking the two syllable homophone 'camel', the name of a kind of animal and the name of a kind of cigarette. In (2) the word 'camel' occurs in a

context which strongly biases expectations towards the cigarette
understanding:

(2) Mary had been smoking for many years. Although she gave
up when she developed bronchitis she still can't resist the
occasional /k æ m.../

The word-initial cohort on the basis of /k æ m.../ might include the
following: cam, camber, camel 1, camel 2, cameo, camera, camisole,
camouflage, camp. campus, kamikaze. I'm not especially concerned with
the possible effects of syntactic constraints here but let's say that
they push out any verb candidates in the pool, e.g. the verbal form
of 'camp'. Interpretative/conceptual constraints, however, should
knock off quite a range of these, in particular they should cause
'camel 1' (the animal concept one) to be dropped. But of course we
know from the Swinney results that at the acoustic offset of
ambiguous words, in this case at the offset of the lateral /l/, both
items are still highly activated, i.e. even given this biasing context
we would find reactions to both of, say, 'desert' and 'cigarette'
significantly facilitated relative to a word unrelated to either
sense. An attempt to marry the cohort view with the Swinney
findings would have to say that 'camel 1' gets dropped from the
cohort on the basis of contextual constraints but somehow
reintroduced when the full input stimulus has been received. This
would seem to be an entirely pointless procedure and it is quite
unclear how it would work.

In fact this is just a special case of a much more general
problem, pointed out by Norris (1982), which is that of accounting
for how it is that contextually inappropriate words are ever
recognised. Pragmatically anomalous uses of words as in (3a) and
even semantically anomalous uses as in (3b) are recognised rapidly
and accurately:

(3) a. John buried the guitar.
b. John drank the guitar.

Marslen-Wilson (1987) himself points this out and it is one of the
considerations lying behind his revised cohort model. He introduces
various changes, the most interesting from our point of view being
the following: "context plays no role in the processes of form-based
access and selection" (1987, 71), "the activation levels (of the
various candidates in the cohort) are a measure of the relative
goodness of fit of the candidates to the bottom-up input, and context
does not tamper with this measure" (1987, 98). In this revised
picture there is no suppressing of one of the 'camel' words by
contextual constraints and it makes perfect sense that there should
be priming of words related to both senses when they are presented
at the offset of the homophone. Context comes into play
subsequently, at the integration phase, and accounts for the rapid
disappearance of equivalent priming. As Marslen-Wilson sees it, both
items are presented in parallel to compete for the available sites in
the higher-level (conceptual) representation. Any decent pragmatic

measure of good fit, such as Sperber & Wilson's principle of relevance for example, will choose 'camel 2' immediately.

All of this seems to comport very nicely with the Fodorian modularity picture. We have bottom-up, encapsulated access and selection processes which are internal to the language module, and these "send up", as it were, a set of candidates to the higher level representation, the message or discourse level, in central systems conceptual format, where they are integrated in accordance with pragmatic principles that operate at that level. If this were all a person knew of Marslen-Wilson's work she might think that what she was encountering was work within a modular framework. Selected passages of Marslen-Wilson and Tyler (1987) (hereafter M-W & T) might confirm this, for example: "top-down influences do not control its (the language processing system's) normal first-pass operations" (M-W & T, 1987, 37). Assuming the following distinction, due to Crain & Steedman (1985):

Weak interactionism: linguistic (syntactic and lexical) processing independently sets up the alternatives which semantics and context choose amongst.

Strong interactionism: semantics and context are predictive, that is, they influence which syntactic/lexical entities are accessed in the first place.

It would seem that both Fodor and M-W & T are maintaining a position of weak interactionism.

3. Strong interactionism and parsing.

The conclusion arrived at in the last section does not however mesh with M-W & T's intentions. The title of their paper is "Against modularity"; in it they challenge Fodor's thesis point by point and they take a strong interactionist position on the processing of syntactic ambiguities, that is, the position that a strongly biased context will actually predict an analysis and thereby cause only one structure to be accessed or computed.

Let's look now at an experiment that they did some time ago on the processing of ambiguous phrases such as 'visiting relatives', 'growing flowers', etc, ambiguous between a verbal/gerund understanding and an adjectival understanding. I'm interested in looking at this experiment, as opposed to any of the others in the 1987 paper, because it seems, at least initially, to give the best support for a strong interactionist position. The structurally ambiguous fragments were presented to the subjects auditorily, through head-phones, the phrase preceded in each case by a clause whose sense should bias expectations towards one of the possible structures. Some examples are given in (4):

(4) Verbal:
          If you want a cheap holiday, visiting relatives ...
          As a traditional way of gaining votes, shaking hands ...
          If you've been trained as a pilot, landing planes ...
     Adjectival:
          If you have a spare bedroom, visiting relatives ...
          If you're trying to thread a needle, shaking hands ...
          If you walk too near a runway, landing planes ...

In both types of case, at the acoustic offset of 'relatives', 'planes',
etc, either IS or ARE was flashed up on a screen and the subjects'
task was to name the word on the screen as quickly as possible. IS
is an appropriate continuation for the verbal case, ARE for the
adjectival, where this appropriateness or otherwise is determined by
the preceding context clause. So, in some cases the word flashed up
was an appropriate continuation, in others inappropriate.
          The results were clearcut and not at all surprising: reaction
times to appropriate probes were significantly faster than reaction
times to inappropriate ones across both the verbal and the adjectival
cases. From this result alone I don't think K-V & T can make any
claims about contextual information (contained in the preceding
clause) directing syntactic analysis, since obviously there is nothing
to stop Fodor or any modularist from claiming that the time
difference is caused by the difficulty of contextually integrating the
inappropriate continuation. That is, the effect is occurring at the
stage of assimilation into the higher level representation rather
than at the stage of accessing of the syntactic structure. Hence the
process affected is not the parse itself but a module-external one.
What's crucial for this to even begin to look like counterevidence to
Fodor is the results that K-V & T got for comparable UNambiguous
phrases, such as those given in (5):

(5) Verbal:
          making movies, mixing drinks, cleaning teeth,
          whistling tunes, firing employees, turning corners,
          carving meat, shouting insults,
     Adjectival:
          creaking stairs, working mothers, travelling
          salesmen, flattering remarks, shooting stars,
          wading pools, dancing classes, landing lights,

These were tested in the same way as outlined above, each being
followed by an appropriate or inappropriate word: IS or ARE
sometimes, but also a range of other verbs, e.g SEEMS or SEEN,
BECOMES or BECOME, etc. Again there was a significant difference in
naming times for the appropriate and inappropriate probes, and it
was pretty much the same sort of difference as for the ambiguous
cases, that is around 30 msecs.
          It is this that gives some force to K-V & T's claims.
According to them, in these cases no decision has to be made between
two syntactic structures, since there is only one possible analysis
in each case. Assuming this, the absolutely parallel results for the

two types of case (ambiguous and unambiguous) can be explained as
follows: in the ambiguous examples it is not the case that both
structures are accessed (computed) with a subsequent choice made
between them; rather the information given in the context clause
which biases for one structure over the other actually feeds directly
into the language processing system and influences the parsing
process so that only one structure is accessed and no process of
disambiguation is required. "Instead of arguing that both analyses
were computed and that one was later selected, we argued that context
affected the parsing process directly, so that only one reading was
ever computed." (M-V & T, 1987, 49, referring to their original
presentation of the experiment in Tyler & Marslen-Wilson, 1977). The
experiment was rerun in 1984 by Marslen-Wilson & Young, to control
for a possible confounding factor (the presence of singular and
plural cataphoric pronouns in the context clause), with essentially
the same pattern of results. We seem then to have here a case of
strong interaction, a case of contextual information, given by the
preceding clause, having a predictive role rather than merely an
accept/reject role.

If this is all correct, selecting a structure when faced with a
syntactic ambiguity is a very different kind of process from the
accessing and selection of a word when faced with a lexical
ambiguity. While lexical selection is entirely bottom-up, stimulus-
driven, etc, syntactic selection can be top-down, directed and
controlled by contextual factors of sense and plausibility. This may
be the way things are, though it does raise a number of questions,
such as (i) Why should such a difference exist? (ii) How is this
strong interactionism possible? That is, how does a higher level
conceptual representation communicate with the processes of syntactic
analysis? Since I don't think the putative difference really does
exist, I shall leave such questions for others to address (those who
are convinced by Marslen-Wilson & Tyler).

4. Distinguishing the unambiguous from the highly plausible

I want to consider now another way of looking at M-V & T's
results, an interpretation which won't raise these sorts of questions.
Let's look again at the allegedly unambiguous phrases in (5).

Now it is generally true that these are conceptually
unambiguous, by which I mean they receive only one interpretation
across a wide range of distinct contexts. You can't get an adjectival
interpretation for the verbal cases: you don't get teeth that clean or
movies that make, tunes that whistle (though there is of course the
interpretation 'tunes for whistling'), drinks that perform mixing
operations, etc. Similarly, the idea is that you can't get verbal
interpretations for the adjectival cases: you can't frighten novels,
travel salesmen, shoot stars, etc.

Even if we are willing to go along with this view, it certainly
doesn't hold for these -ing words in general, at least not for those
which have transitive verbs as their base. And, as it happens, all
of the examples used by T & M-V (1977) did involve transitive verbs.
In the 1987 paper, however, they mention the example of 'smiling

faces' and 'smile' is presumably subcategorised as intransitive so
that 'smiling' has an adjectival relation to the following noun.
Perhaps this example was used in the unpublished Marslen-Wilson &
Young rerun of the experiment, but, even if it was, the overwhelming
majority of verbs used were transitive. This means that all the -ing
words taken alone could be interpreted as either gerunds or
adjectives, and since most of them can take objects the V-ing + NP
configuration also has the two possible syntactic analyses.

What this suggests to me is that actually all these phrases are
structurally/syntactically ambiguous, and that it's the semantic/
conceptual content of the phrase as a whole which determines whether
the -ing word is verbal or adjectival. Given that, it is by no means
obvious that only one syntactic structure is accessed even in these
supposedly unambiguous cases. Rather, I would say a verbal and an
adjectival reading for each phrase is accessed and sent on up to the
central systems for contextual integration. Here the intended
(plausible/relevant/sensible) meaning is established via interaction
with real world non-linguistic knowledge about the relationship of
cleaning and teeth, turning and corners, mixing and drinks,
frightening and novels, shooting and stars, etc.

Such an account is a complete reversal of M-W & T's analysis.
They start from the assumption that only one structure is accessed
by the language input system in their "unambiguous" cases and, since
reaction times for the ambiguous cases are much the same as for the
"unambiguous" cases, their conclusion is that here too it must be that
only one structure is accessed. I'm starting from an assumption that
two structures are accessed in the ambiguous cases and arguing that
the same goes for the so-called unambiguous cases because they are
in fact structurally ambiguous. Certainly, in most of the cases in
(5) the plausibility of one of the possible interpretations is
overwhelmingly greater than the other but there is no reason to
suppose that this conceptual disambiguation is a job performed by the
processes responsible for syntactic analysis, dependent as it is on
encyclopaedic information.

In fact we wouldn't want such possibilities as a gerund
interpretation of 'shooting stars' (e.g. 'shooting stars is my
favourite hobby'), or an adjectival interpretation of 'whistling tunes'
(tunes good for whistling), to be ruled out as impossible analyses.
Even some of the weirder ones, such as teeth that clean, might occur
in certain fictional contexts. It would be hopeless if these were
treated uniformly as syntactically unambiguous; it would mean we
could never understand them, which is simply not the case, just as we
can understand lexical items even when their use is semantically or
pragmatically anomalous.

In all these cases then, whether conceptually ambiguous (two
plausible readings) or unambiguous (one much more plausible than the
other across the majority of contexts), two syntactic structures may
be accessed and presented in parallel to the higher level discourse
representation which makes a choice between them on the basis of
background and contextual information. In Marslen-Wilson's words we
thereby "exploit the capacity of a parallel system for multiple
access and multiple assessment" (1987, 100), i.e. the central

integrative system can pursue several lines of analysis simultaneously. If this view of the processing of syntactic ambiguities is correct, what looked initially like a case of strong interactionism turns out instead to be weak. It is not, after all, radically different in this respect from the processes of word recognition but, like them, is quite consistent with the modularity of the language system.

5. Evidence for a parsing strategy?

I wouldn't in fact want to cling adamantly to the parallel accessing and processing of these structures. There may be a preferred structure which is always tried first, with the second one considered only if the first one is rejected by the contextual integration processes. There have been various proposals along these lines for parsing strategies which achieve overall economy and speed in processing at the price of sometimes making the wrong first shot. So, for instance, if Lyn Frazier's Minimal Attachment Strategy is a heuristic employed by the language input system the prediction is that in many cases of syntactic ambiguity one structure is given first try-out. This strategy is given in (6):

(6) Incorporate each item into the evolving structure using the fewest nodes consistent with the grammar.
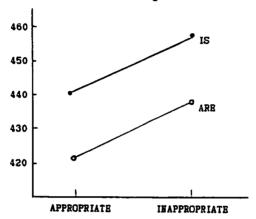
The existence of such strategies is quite consonant with the modularity of the language system: Frazier's idea is that they are part of the language module and thus take priority over any broader contextual influence on parsing, which may however over-rule them at a later stage in processing (see Frazier, 1985, and Ferreira and Clifton, 1986, for evidence that this is so).

What this particular strategy would predict for the sort of phrases we've been considering, or indeed for the -ing word alone, is that the adjectival analysis is chosen first since a verbal analysis would require more nodes, presumably at least an additional S-node and an empty category NP subject. The system would always try the adjectival analysis first; in the transitive cases this would be confirmed by the following noun and the syntactic analysis would continue smoothly. Subsequent submission to the higher level discourse representation naturally brings into play constraints of sense and plausibility. If the analysis proves unsuccessful here, it is rejected and the parser has to try again, in accordance with the sort of weak interaction between parser and semantic context that Fodor envisages. In this case the parser takes up the other option offered by the grammar, the option which isn't in accord with the parsing strategy, and submits it for contextual integration.

Why should we bother to consider this possibility? After all, parallel accessing and assessment are much more in keeping with the current view of the human mind as capable of a massive amount of parallel computing and certainly the evidence seems to be that lexical access, selection and integration works like this. Well, in addition to the evidence that Frazier and her colleagues have for the

strategy (which I won't review here; see, for example, Frazier, 1987, and Clifton & Ferreira, 1987) it looks as if it might explain something puzzling in M-V & T (1987). When they talk briefly there about the rerun of the syntactic ambiguity experiment done by Marslen-Wilson & Young they summarise the results in a graph, which is given in (7):

(7)   Mean naming latencies (in msecs) for appropriate and inappropriate IS and ARE targets:

We see here the significant effect of contextual appropriateness on response times to the IS and ARE probes, a difference of around 20 msecs. in each case, but what is curious is that the two lines don't coincide. Why was the naming of IS always slower than the naming of ARE, even when IS was an appropriate continuation and ARE was an inappropriate one? M-V & T don't remark on this.

Serial access of syntactic structures in accordance with some preferential strategy such as the Minimal Attachment Strategy would provide an explanation. If, when confronted with one of these -ing phrases, the parser goes for the adjectival reading, then, given the plural noun in all the examples here, the continuation of ARE is, at this stage, more appropriate than IS. Only when the structure proves unable to integrate with the preceding context clause and so is rejected will the gerund structure be accessed. If this is the way things work, naming of IS could never be faster than naming of ARE and that's exactly what the graph shows us. This then looks like suggestive evidence for such a strategy, imposing serial access.

Either way, whether there is multiple access from the beginning or serial access in accordance with some parsing strategy, we have a case of a linguistic process which interacts weakly with context and which bears all the marks of a modular process.

6. Concluding comments

If cognitive processors are computational systems, and the modularity dispute considered here takes this assumption to be common ground, then "they have access to information solely in virtue of the _form_ of the representations in which it is couched" (Fodor, 1983, 40-41). That is, computational processes are _syntactic_. This gives rise to what can be called "the format constraint" on mental processing: neither acoustic representations nor primal sketches, for instance, can interact inferentially with conceptual thought, nor, presumably with each other. As Fodor (1983, 40) puts it: "what perception must do is to so represent the world as to make it accessible to thought. The condition on appropriateness of format is by way of emphasising that not every representation of the world will do for this purpose". It is the systems responsible for achieving commonality of format, that is, the perceptual systems and language, that Fodor proposes as mental modules.

This view of cognitive architecture as in part modular and in part nonmodular (the central conceptual systems) seems to offer just the right amount of flexibility: it is an interactive picture with only such restrictions on interaction as make it possible to conceive of how the whole system might actually work. Encyclopaedic and contextual information does not direct syntactic processing because it cannot; the two levels of representation are not yet in a common language enabling them to communicate. If the format constraint on mental processing is a necessary constraint on processing, and it seems to be so, at least within a symbolic computational theory of mind, then Fodorian modularity gives us an account of how information from a range of sources, auditory, visual and linguistic, together with existing beliefs, can ultimately get together in a common code and interact inferentially in the ongoing process of establishing an accurate representation of the world.

7. References

Carston, R. (1988) "Language and cognition". In Newmeyer, F. (ed) Linguistics: The Cambridge Survey Vol III. Cambridge University Press.

Clifton, C. & Ferreira, F. (1987) "Modularity in sentence comprehension". In Garfield, J. (ed) Modularity in Knowledge Representation and Natural Language Understanding. MIT Press.

Crain, S. & Steedman, M. (1985) "On not being led up the garden path: the use of context by the psychological syntax processor". In Dowty, D., Karttunen, L. & Zwicky, A. (eds) Natural Language Parsing. Cambridge University Press.

Fodor, J.A. (1983) Modularity of Mind. MIT Press.

Fodor, J.A. (1986) "The modularity of mind". Chapter 1 of Pylyshyn, Z.W. & Demopoulos, W. (eds) Meaning and Cognitive Structure. Ablex.

Ferreira, F. & Clifton, C. (1986) "The independence of syntactic processing". Journal of Memory and Language 25, 348-368.

Frazier, L. (1985) "Modularity and the representational hypothesis". Proceedings of the Northeastern Linguistics Society, 131-144.

Frazier, L. (1987) "Theories of sentence processing". In Garfield, J. (ed) Modularity in Knowledge Representation and Natural Language Understanding. MIT Press.

Marslen-Wilson, W. (1987) "Functional parallelism in spoken word-recognition". Cognition 25, 71-102.

Marslen-Wilson, W. & Tyler, L. (1981) "Central processes in speech understanding", Philosophical Transactions of the Royal Society B 295, 317-332.

Marslen-Wilson, W. & Tyler, L. (1987) "Against modularity".   In Garfield, J. (ed) Modularity in Knowledge Representation and Natural-Language Understanding. MIT Press.

Norris, D. (1982) "Autonomous processes in comprehension", Cognition 11, 97-101.

Onifer, W. & Swinney, D. (1981) "Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias", Memory & Cognition 9 (3), 225-236.

Seidenberg, M., Tanenhaus, M., Leiman, J. & Bienkowski, M. (1982) "Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing", Cognitive Psychology 14, 489-537.

Sperber, D. & Wilson, D. (1986) Relevance: Communication and Cognition. Blackwell.

Swinney, D. (1979) "Lexical access during sentence comprehension: (Re)consideration of context effects", Journal of Verbal Learning and Verbal Behavior, 18, 645-660.

Tanenhaus, M. & Lucas, M. (1987) "Context effects in lexical processing", Cognition 25, 213-34.

Tyler, L. & Marslen-Wilson, W. (1977) "The on-line effects of semantic context on syntactic processing", Journal of Verbal Learning and Verbal Behavior 16, 683-692.